

# Registrierung und Segmentierung von Lymphknoten aus multimodalen Zeitreihen im Kopf-Hals-Bereich



vom Fachbereich Informatik  
der Technischen Universität Darmstadt  
genehmigte

## DISSERTATION

zur Erlangung des akademischen Grades eines  
Doktor-Ingenieurs (Dr.-Ing.)  
von

Dipl.-Ing. Sebastian Steger  
geboren in München, Deutschland

Referenten der Arbeit: Prof. Dr.-Ing. Georgios Sakas  
Technische Universität Darmstadt

Prof. Dr. techn. Dieter W. Fellner  
Technische Universität Darmstadt

Tag der Einreichung: 19.12.2012  
Tag der mündlichen Prüfung: 07.02.2013

Darmstädter Dissertation  
D 17

Darmstadt, 2013



# Erklärung zur Dissertation

Hiermit versichere ich die vorliegende Dissertation selbständig nur mit den angegebenen Quellen und Hilfsmitteln angefertigt zu haben. Alle Stellen, die aus Quellen entnommen wurden, sind als solche kenntlich gemacht. Diese Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

Darmstadt, den 19.12.2012

Sebastian Steger





Für Sohan



# Danksagung

An dieser Stelle möchte ich mich bei einer ganzen Reihe von Personen herzlich bedanken, die zum Gelingen dieser Arbeit beigetragen haben. An erster Stelle danke ich Prof. Dr.-Ing. Georgios Sakas für die Möglichkeit, die Promotion durchzuführen und für seine uneingeschränkte Unterstützung während des gesamten Promotionsprozesses. Als mein direkter Vorgesetzter am Fraunhofer IGD ließ er mir die Freiheit meine eigenen Ideen zu verwirklichen.

Dem Koreferenten dieser Arbeit, Prof. Dr. techn. Dieter W. Fellner, möchte ich ebenfalls meinen Dank aussprechen. Seine kritische Betrachtung der Arbeit aus Sicht der Informatik losgelöst von der medizinischen Fragestellung ermöglichte mir die Arbeit zu verbessern.

Auch Dr. Stefan Wesarg, der als Nachfolger von Prof. Sakas als Abteilungsleiter mir die Möglichkeit gegeben hat, meine Arbeit fortzusetzen, gilt mein Dank.

Weiterhin möchte ich mich bei dem Radiologenteam von Dr. Gianfranco Chiari an dem Universitätsklinikum in Parma, Italien, bedanken. Dort bekam ich einen umfangreichen Einblick in die klinische Vorgehensweise und deren Limitierungen. Dr. Chiari hat mir auf leicht verständliche Weise ein umfangreiches medizinisches Wissen vermittelt und mich bei meiner Arbeit aus klinischer Sicht unterstützt. Insbesondere habe ich von ihm umfangreiche Testdaten erhalten, die er mit seinem Team manuell annotierte. Ohne diese Daten hätte ich weder die Entwicklung noch die Evaluation durchführen können.

Auch meinen Kollegen möchte ich in diesem Zuge für die vielen fachlichen Diskussionen, kritischen Kommentare und die ausgezeichnete Arbeitsatmosphäre meinen Dank aussprechen. Besonders hervorheben möchte ich dabei Matthias Kirschner, dessen Softwarebibliothek ich verwenden konnte. Für ihre Beiträge danke ich auch meinen ehemaligen Studenten Nazli Bozoglu und Daniel Ebert.

Für das Korrekturlesen der Arbeit möchte ich mich bei Evelyne Neuberger, Dr. Stefan Wesarg und Seena Vellaramkalayil bedanken.

Ein ganz persönlicher Dank gilt meiner wunderbaren Ehefrau Teena. Sie hat mich nicht nur durch ihr volles Verständnis für die vielen Abende und Wochenenden, die ich mit dem Schreiben dieser Arbeit verbracht habe, unterstützt, sondern auch durch die vielen zielführenden fachlichen Diskussionen, die zu einer Verbesserung dieser Arbeit beitrugen. Nach der Geburt unseres Sohnes in den Endzügen der Arbeit hat sie mir auf großartige Weise den Rücken freigehalten und mich in den schlaflosen Nächten vor der Verteidigung aufgebaut.



# Kurzfassung

Der verlässlichste unabhängige prognostische Faktor für den Krankheitsverlauf von Patienten mit Kopf-Hals-Karzinom ist das Vorhandensein von Lymphknotenmetastasen. Eine computergestützte Untersuchung und zeitliche Verfolgung von Lymphknoten in mehreren Bildmodalitäten durch ein multimodales, multitemporales Modell bietet viele Vorteile, insbesondere in Bezug auf Reproduzierbarkeit. Grundvoraussetzung dafür ist jedoch eine robuste automatische Registrierung und Segmentierung von Lymphknoten aus multimodalen Zeitreihen. Da existierende Verfahren den Anforderungen nicht genügen, werden dazu in dieser Arbeit neuartige Methoden entwickelt und evaluiert.

Zur Lymphknotensegmentierung aus CT-Datensätzen wird ein Radialstrahl-basiertes 3D-Verfahren umfangreich behandelt. Ausgehend von einem Saatpunkt werden Strahlen radial in alle möglichen Richtungen gleichverteilt geschickt und ein Optimierungsverfahren bestimmt für jeden Strahl unter Einbeziehung von Bildinformation und lokalem Formwissen den bestmöglichen Radius und somit eine Segmentierung. Erstmals findet ein Vergleich unterschiedlicher bildbasierter Kostenfunktionen statt und die Parameter werden durch ein datengetriebenes Verfahren bestimmt. Mit einer durchschnittlichen Oberflächendistanz von nur  $0.46 \text{ mm}$  ist die Segmentierungsgenauigkeit im Bereich der manuellen Expertensegmentierung und deutlich besser als existierende semi-automatische Verfahren. Die Inter-Observer-Variabilität zur Volumenbestimmung ist um den Faktor 3 geringer als bei manueller Volumenbestimmung. Neben Lymphknoten eignet sich das Verfahren auch zur Segmentierung anderer rundlicher Strukturen, wie z.B. Tumore, und bietet auch für einige Organe, wie z.B. die Prostata, eine Alternative zur modellbasierten Segmentierung.

Die Registrierung einzelner Lymphknoten erfordert die automatische deformierbare Registrierung des gesamten Kopf-Hals-Bereichs. Dazu wird in dieser Arbeit die erste vollautomatische generalisierbare multi-rigide Methode vorgestellt. Sie basiert auf einem neuartigen artikulierten Atlas, welcher neben Wissen über Form und Aussehen einzelner Knochen auch deren relative Lagen (Artikulation) aus Trainingsdaten lernt. Dieser wird zunächst zur gleichzeitigen Segmentierung der Knochen aus dem CT-Datensatz verwendet. Ausgehend davon wird er durch personalisiertes Wissen angereichert und an die andere Bildmodalität oder Zeitserienaufnahme unter Berücksichtigung des gelernten Artikulationsraums angepasst. Die daraus berechneten rigiden Transformationen werden in einem zweistufigen Prozess in das umliegende Weichteilgewebe propagiert und ein dichtes Deformationsfeld entsteht. Abschließend wird die Registrierungsgenauigkeit innerhalb der Lymphknoten durch ein lokal rigides Registrierungsverfahren verbessert. Die Vorteile der multi-rigiden Registrierung liegen in dem großen Konvergenzbereich und der geringen Anfälligkeit für Bildartefakte aufgrund der einzigartigen globalen Regularisierung. Innerhalb von Lymphknotenzentren wird eine Registrierungsgenauigkeit von durchschnittlich  $5.05 \text{ mm}$  erreicht. Gegenüber der B-Spline Registrierung ergibt sich eine Verbesserung um 37% und eine subjektiv deutlich besser empfundene Qualität der erzeugten deformierten Bilder.

Die wichtigsten Beiträge dieser Arbeit sind zur Segmentierung von Lymphknoten eine neue Kostenfunktion, ein umfangreicher Vergleich verschiedener Kostenfunktionen und eine datengetriebene Parameterwahl. Die Hauptbeiträge zur Bildregistrierung sind das Lernen der relativen Elementlagen, ein adaptives Anpassungsverfahren und die Personalisierung des artikulierten Atlas.



# Abstract

The most reliable independent prognostic factor for the disease evolution of head&neck cancer patients is the presence of lymph node metastasis. A computer aided evaluation and follow up assessment of relevant lymph nodes in different imaging modalities using a multimodal multitemporal model has many advantages, in particular in terms of reproducibility. However, this requires the robust automated registration and segmentation of lymph nodes from multi-modal time series medical images. In this work, new approaches for both imaging problems are presented because existing solutions do not satisfy the requirements.

For the segmentation of lymph nodes from CT-data, a radial ray based 3D-method is treated comprehensively. From a seed point, rays are cast into all directions and an optimization technique determines a radius for each ray based on image appearance and shape knowledge thereby obtaining a segmentation. For the first time, different appearance cost functions are compared and parameters are selected using a data driven approach. With an average symmetric surface distance of only  $0.46\text{ mm}$ , the segmentation accuracy is close to manual segmentation and outperforms existing semi-automatic methods. The inter-observer-variability for volume assessment is approximately three times lower compared to manual volume assessment. Beyond lymph nodes, this method is also suitable for the segmentation of other spherical structures such as tumors and offers an alternative to model based segmentation approaches for certain organs like the prostate.

Prerequisite for a registration of individual lymph nodes is the automated deformable registration of the entire head&neck images. Therefore, the first fully automated general purpose multi-rigid registration method is presented. It is based on an articulated atlas combining shape and appearance knowledge of individual bones with information about their relative spatial locations (articulation). It is created from annotated training data and used to obtain a segmentation of individual bones from CT-data. Based on that, the atlas is personalized and then adapted to the other imaging modality or time series image, thereby incorporating the learned articulation space. The resulting rigid transformations for each bone are propagated into the surrounding soft tissue in a two step process resulting into a dense deformation field. Finally, the accuracy within the lymph nodes is improved using local rigid registration. Advantages of the multi-rigid image registration approach presented in this work are the huge capture range as well as low sensitivity towards imaging artifacts due to global regularization. The average registration accuracy in the lymph node centers is  $5.05\text{ mm}$ . Compared to B-spline registration, this is 37% better while at the same time a considerable better quality of the deformed images is achieved.

The major contributions of this work to the segmentation of lymph nodes are a new cost function, a comprehensive comparison of cost functions and a data driven parameter selection. The main contributions to image registration are the learned articulation, a dynamic adaptation strategy and the personalization of the articulated atlas.





# Inhaltsverzeichnis

|  |           |
|--|-----------|
| <b>1. Einleitung</b>   | <b>1</b>  |
| 1.1. Motivation . . . . .  | 1         |
| 1.2. Anwendungsszenario . . . . .  | 4         |
| 1.3. Zielsetzung . . . . .   | 5         |
| 1.3.1. Segmentierung von Lymphknoten . . . . .                                       | 5         |
| 1.3.2. Registrierung von Lymphknoten . . . . .                                       | 6         |
| 1.4. Aufbau der Arbeit . . . . .   | 7         |
| <b>2. Grundlagen</b>   | <b>9</b>  |
| 2.1. Lymphknoten . . . . .   | 9         |
| 2.1.1. Eigenschaften . . . . .   | 9         |
| 2.1.2. Lymphadenopathie . . . . .  | 11        |
| 2.1.3. Bildgebung . . . . .  | 13        |
| 2.2. Segmentierung und Registrierung in der medizinischen Bildverarbeitung . . . . . | 17        |
| 2.2.1. Segmentierung . . . . .   | 17        |
| 2.2.2. Registrierung . . . . .   | 20        |
| 2.2.3. Regmentierung . . . . .   | 26        |
| <b>3. Stand der Forschung</b>  | <b>29</b> |
| 3.1. Segmentierung von Lymphknoten aus CT-Datensätzen . . . . .                      | 29        |
| 3.1.1. Lageorientierte Verfahren zur Lymphknotenregionensegmentierung . . . . .      | 29        |
| 3.1.2. Vergleichende Studien . . . . .   | 30        |
| 3.1.3. Grauwert-, Regionen- und Kanten-orientierte Verfahren . . . . .               | 32        |
| 3.1.4. Verfahren mit globalem Formwissen . . . . .                                   | 32        |
| 3.1.5. Verfahren mit lokalem Formwissen . . . . .                                    | 37        |
| 3.1.6. Zusammenfassung . . . . .   | 42        |
| 3.2. Registrierung des Kopf-Hals-Bereichs . . . . .                                  | 46        |
| 3.2.1. Transformationsklassen . . . . .  | 47        |
| 3.2.2. Bestimmung der einzelnen Rigid Strukturen . . . . .                           | 50        |
| 3.2.3. Registrierung der rigid Strukturen . . . . .                                  | 55        |
| 3.2.4. Bestimmung des Deformationsfeldes im Weichteilgewebe . . . . .                | 56        |
| 3.2.5. Zusammenfassung . . . . .   | 58        |
| <b>4. Methoden</b>   | <b>61</b> |
| 4.1. Segmentierung von Lymphknoten in CT-Datensätzen . . . . .                       | 61        |
| 4.1.1. Größeninvariantes Masse-Feder-Modell . . . . .                                | 61        |
| 4.1.2. Radialstrahl-basierte 2D-Segmentierung von Lymphknoten . . . . .              | 64        |
| 4.1.3. Radialstrahl-basierte 3D-Segmentierung von Lymphknoten . . . . .              | 67        |

|           |  |            |
|-----------|--|------------|
| 4.1.4.    | Optimierungsverfahren und Formerhaltung . . . . .  | 68         |
| 4.1.5.    | Bildgradientenbasierte Kostenfunktionen . . . . .  | 72         |
| 4.1.6.    | Intensitätsbasierte Kostenfunktionen . . . . .   | 78         |
| 4.1.7.    | Parameterbestimmung . . . . .  | 83         |
| 4.1.8.    | Erweiterungen . . . . .  | 92         |
| 4.1.9.    | Zusammenfassung . . . . .  | 98         |
| 4.2.      | Registrierung von Lymphknoten in Kopf-Hals-Tomographien . . . . .                          | 101        |
| 4.2.1.    | Multimodale rigide Registrierung des Kopfes . . . . .                                      | 101        |
| 4.2.2.    | Artikulierter Atlas zur Segmentierung einzelner Knochen . . . . .                          | 107        |
| 4.2.3.    | Anpassung des artikulierten Atlas . . . . .  | 115        |
| 4.2.4.    | Sukzessive lokal rigide Registrierung einzelner Knochen . . . . .                          | 122        |
| 4.2.5.    | Personalisierter artikulierter Atlas . . . . .   | 124        |
| 4.2.6.    | Weichteilgewebe-Deformation . . . . .  | 125        |
| 4.2.7.    | Weitere Stützpunkte . . . . .  | 128        |
| 4.2.8.    | Lokal rigide Registrierung von Lymphknoten . . . . .                                       | 131        |
| 4.2.9.    | Zusammenfassung . . . . .  | 132        |
| <b>5.</b> | <b>Experimente &amp; Ergebnisse</b>  | <b>135</b> |
| 5.1.      | Segmentierung von Lymphknoten . . . . .  | 135        |
| 5.1.1.    | Material . . . . .   | 135        |
| 5.1.2.    | Größeninvariantes Masse-Feder-Modell . . . . .   | 138        |
| 5.1.3.    | Segmentierung von Lymphknoten in einzelnen CT-Schichten . . . . .                          | 140        |
| 5.1.4.    | Radialstrahl-basierte Segmentierung . . . . .  | 141        |
| 5.1.5.    | Zusammenfassung . . . . .  | 157        |
| 5.2.      | Registrierung von Lymphknoten in Kopf-Hals-Tomographien . . . . .                          | 159        |
| 5.2.1.    | Material . . . . .   | 159        |
| 5.2.2.    | Multimodale rigide Registrierung des Kopfes . . . . .                                      | 159        |
| 5.2.3.    | Artikulierter Atlas zur Segmentierung einzelner Knochen . . . . .                          | 161        |
| 5.2.4.    | Registrierung einzelner Knochen . . . . .  | 164        |
| 5.2.5.    | Weichteilgewebe-Deformation . . . . .  | 167        |
| 5.2.6.    | Lokal rigide Registrierung von Lymphknoten . . . . .                                       | 174        |
| 5.2.7.    | Laufzeit . . . . .   | 174        |
| 5.2.8.    | Zusammenfassung . . . . .  | 175        |
| <b>6.</b> | <b>Anwendungen</b>   | <b>179</b> |
| 6.1.      | Lymphknotenanalyse einer multimodalen Zeitreihe zur Krankheitsverlaufsvorhersage . . . . . | 179        |
| 6.2.      | Weitere Anwendungen . . . . .  | 184        |
| 6.2.1.    | Segmentierung von Lebertumoren . . . . .   | 184        |
| 6.2.2.    | Segmentierung der Prostata aus multiparametrischen MRT-Aufnahmen . . . . .                 | 184        |
| 6.2.3.    | Lymphoma-Therapieerfolgsbestimmung in PET-CT-Zeitserienaufnahmen . . . . .                 | 185        |
| 6.2.4.    | Multi-rigide Registrierung der Leber . . . . .   | 186        |
| 6.2.5.    | Segmentierung von Knochen in MRT-Aufnahmen . . . . .                                       | 187        |
| 6.2.6.    | Entfernung von Markern aus Fluoroskopieaufnahmen . . . . .                                 | 188        |
| 6.3.      | Denkbare Anwendungsgebiete . . . . .   | 189        |
| 6.3.1.    | Segmentierung rundlicher Strukturen . . . . .  | 189        |
| 6.3.2.    | Gemeinsame Segmentierung mehrerer Strukturen durch den artikulierten Atlas . . . . .       | 189        |

|  |            |
|--|------------|
| 6.3.3. Deformierbare Registrierung auf Grundlage rigider Strukturen . . . . .                                      | 190        |
| <b>7. Zusammenfassung &amp; Ausblick</b>   | <b>191</b> |
| 7.1. Radialstrahl-basierte Segmentierung rundlicher Strukturen . . . . .   | 191        |
| 7.2. Bildregistrierung auf Grundlage rigider Strukturen . . . . .  | 194        |
| 7.3. Konzeptioneller Vergleich der Radialstrahl-basierten Segmentierung mit dem artiku-<br>lierten Atlas . . . . . | 198        |
| 7.4. Weiterführende Arbeit . . . . .   | 199        |
| <b>Eigene Literatur</b>  | <b>201</b> |
| <b>Betreute Abschlussarbeiten</b>  | <b>203</b> |
| <b>Literaturverzeichnis</b>  | <b>205</b> |
| <b>Abkürzungsverzeichnis</b>   | <b>217</b> |
| <b>Lebenslauf</b>  | <b>219</b> |



# 1. Einleitung

## 1.1. Motivation

Die Untersuchung von Lymphknoten im Kopf-Hals-Bereich mittels bildgebender Verfahren unterschiedlicher Bildmodalitäten hat eine große Bedeutung bei der Stadienbestimmung und Verlaufskontrolle von Patienten mit Kopf-Hals-Karzinom. Im Vergleich zu einer manuellen Untersuchung unterliegt eine computergestützte Auswertung nicht der Subjektivität eines Radiologen und ist daher wünschenswert. Allerdings ist dazu eine Registrierung der herangezogenen Bildmodalitäten sowie die Segmentierung einzelner relevanter Lymphknoten erforderlich. Diese Arbeit widmet sich diesen beiden bisher unzureichend betrachteten Aspekten der medizinischen Bildverarbeitung, indem neuartige Lösungen dafür entwickelt und evaluiert werden.

**Klinische Problemstellung:** An Kopf-Hals-Karzinomen, also Krebserkrankungen in der Mundhöhle, dem Rachen und dem Kehlkopf, erkranken weltweit ca. 550000 Menschen pro Jahr. Damit ist diese Krebserkrankung die siebt-häufigste [19] und die Fünfjahresüberlebensrate liegt in den westlichen Ländern bei nur knapp 40% [20]. Darüber hinaus beeinträchtigen Kopf-Hals-Karzinome oft die Lebensqualität erheblich, insbesondere die Fähigkeit zu sprechen und zu schlucken. Bei der großen Mehrheit der Tumore handelt es sich um das Plattenepithelkarzinom, welches oftmals von regionalen Metastasen in den Lymphknoten (siehe Abbildung 1.1(b)) begleitet wird.

Das Vorhandensein von Lymphknotenmetastasen ist der verlässlichste unabhängige prognostische Faktor für den Krankheitsverlauf [21]. Nach erfolgreicher Behandlung, welche meist durch einen chirurgischen Eingriff erfolgt und oftmals von Strahlen- und seltener auch Chemotherapie begleitet wird, ist die Rezidivquote in den ersten fünf Jahren mit 25 – 50% beachtlich hoch. Da für viele der Patienten das Auftreten eines Rezidivs einem Todesurteil gleichkommt, ist die Beobachtung des Kopf-Hals-Bereichs insbesondere in den ersten zwei Jahren nach der Behandlung von großer Bedeutung.

**Bedeutung bildgebender Verfahren:** Bildgebende Verfahren spielen bei der Diagnose nur eine untergeordnete Rolle, weil der Tumor meist an einer Schleimhaut entdeckt wird und so eine histologische Diagnose auf Basis einer Biopsieprobe oft bereits vor der Bildakquisition vorliegt. Daher liefern bildgebende Verfahren lediglich Zusatzinformationen über Größe des Primärtumors und dessen Lage zu umliegenden Gewebe sowie über das Vorhandensein von regionalen und Fernmetastasen. Bei der Stadienbestimmung, der Therapieplanung und -verlaufskontrolle sowie der Nachverfolgung haben bildgebende Verfahren jedoch eine tragende Rolle [22]. Kontrastmittelverstärktes CT sowie T1- und T2-gewichtetes MRT werden in der klinischen Routine zur Lokalisation und Größenbestimmung des Primärtumors und der Untersuchung von regionalen Metastasen eingesetzt. Trotz der hohen Sensitivität kommt PET aufgrund geringer Spezifität und der hohen Kosten nicht so oft zum Einsatz.

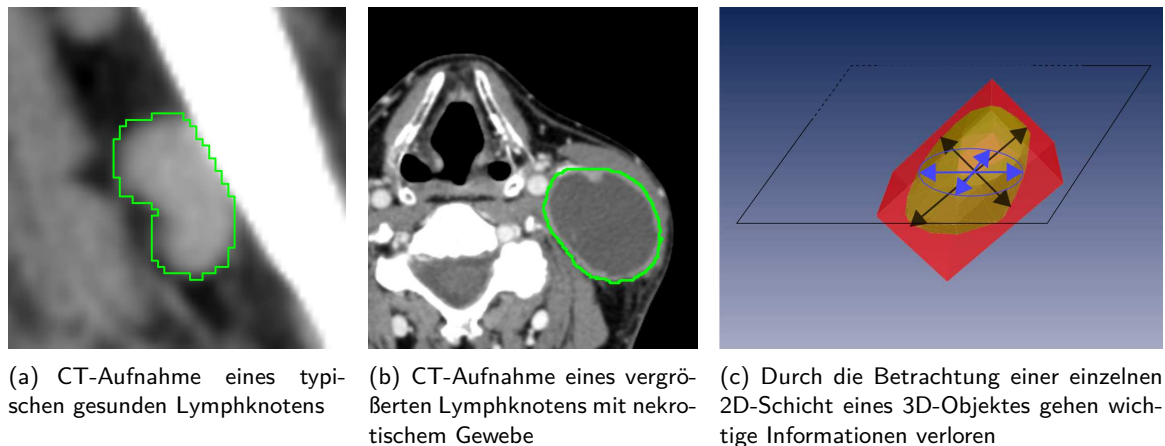


Abbildung 1.1.: Vergleich zwischen 2D- und 3D-Betrachtung eines Lymphknotens

Ultraschall kann lediglich zur Navigation einer Feinnadelbiopsie und als kostengünstige Alternative in der Nachuntersuchung verwendet werden.

Die Stadienbestimmung hat einen großen Einfluss auf die Prognose des Krankheitsverlaufs und bestimmt maßgeblich die Art der eingesetzten Therapie. Sie geschieht international standardisiert nach der TNM-Klassifikation [23], welche auf Größe und Ausbreitung des Primärtumors, den Befall der regionalen Lymphknoten und dem Vorhandensein von Fernmetastasen basiert. Der Lymphknotenbefall wird üblicherweise durch eine radiologische Befundung festgestellt, weil diese eine bessere Genauigkeit als eine klinische Untersuchung durch Abtastung liefert. Dabei ist es wichtig, pathologische von nicht pathologischen Lymphknoten zu unterscheiden. Indikatoren für Pathologien sind unter anderem Vergrößerung, eine rundere Form, Cluster-Bildung, selektive Kontrastmittelaufnahme, Inhomogenität und das Vorhandensein von nekrotischem Gewebe oder einer extrakapsulären Ausbreitung [22][21]. Da einige dieser Indikatoren besser auf CT-Aufnahmen, andere jedoch besser auf MRT-Aufnahmen bestimmt werden können, bietet eine gemeinsame Auswertung beider Bildmodalitäten einen Mehrwert, weshalb oftmals Aufnahmen beider Modalitäten erstellt werden.

**Größenbestimmung von Läsionen:** In der klinischen Praxis erfolgt die Größenbestimmung des Primärtumors und der Lymphknoten üblicherweise mithilfe der im Jahr 2000 vorgestellten RECIST-Kriterien [24] bzw. deren Nachfolger aus dem Jahr 2009 [25][26]. Diese messen allerdings nicht die wahre Größe, also das Volumen, sondern lediglich die 2D-Abmessungen in einer repräsentativen axialen Schicht. Dies hat zwei entscheidende Nachteile. Zum einen sind die Messungen nicht robust reproduzierbar, es besteht also eine hohe Inter- und Intra-Observer-Variabilität. Zum anderen geht prinzipbedingt die 3D-Information durch die Betrachtung einer einzelnen axialen Schicht verloren, was zur Folge hat, dass diese Messungen nur unzureichend mit dem tatsächlichen Volumen korrelieren (siehe Abbildung 1.1(c)). Trotz dieser Nachteile findet in der klinischen Praxis eine volumetrische Betrachtung nicht statt, weil eine 3D-Modellerstellung durch manuelle Konturierung zu zeitaufwändig ist. Die manuelle schichtweise Konturierung eines einzelnen Lymphknotens nimmt meinen eigenen Messungen in [1] bis zu 25 Minuten in Anspruch. (Semi-) automatische Verfahren zur 3D-Segmentierung von Läsionen, die diese Nachteile nicht haben, werden daher benötigt.

| Bildverarbeitungsmethode                | Anwendung                                  |
|---|--|
| Segmentierung des Primärtumors          | Stadienbestimmung, Strahlentherapieplanung |
| Segmentierung von Lymphknoten           | Stadienbestimmung, Verlaufskontrolle       |
| Detektion von Lymphknoten               | Stadienbestimmung, Verlaufskontrolle       |
| Segmentierung der Lymphknotenregionen   | Strahlentherapieplanung                    |
| Registrierung von einzelnen Lymphknoten | Stadienbestimmung, Verlaufskontrolle       |
| Reg. des gesamten Kopf-Hals-Bereichs    | Verlaufskontrolle, Strahlentherapieplanung |

Tabelle 1.1.: Benötigte Bildverarbeitungsmethoden zur Behandlung von Patienten mit Kopf-Hals-Karzinomen

**Strahlentherapieplanung:** Zusätzlich zu der chirurgischen Entfernung des Primärtumors und der befallenen regionalen Lymphknoten wird bei der Behandlung von Kopf-Hals-Karzinomen oftmals auch die Strahlentherapie eingesetzt. Dabei soll so viel zerstörerische Strahlung wie möglich in die Zielregionen und so wenig wie möglich in nicht betroffenes Gewebe, insbesondere in die Risikoorgane wie z.B. den optischen Nerv, die Speicheldrüsen, das Rückenmark oder den Hirnstamm, gelangen. Dies erfordert eine präzise Konturierung der entsprechenden Regionen auf einem sog. Planungs-CT, was sehr zeitaufwändig ist und einer hohen Inter-Observer-Variabilität unterliegt. Letzteres kann durch die zusätzliche Einbeziehung komplementärer Bildaufnahmetechniken, wie z.B. MRT oder PET, verringert werden [21]. Sowohl das Übertragen der Planung in die Therapie als auch die Einbeziehung mehrerer Modalitäten bei der Konturierung erfordert allerdings, dass die örtlichen Zusammenhänge bekannt sind. In der Praxis ist dies meistens dadurch sichergestellt, dass dem Patienten eine Maske aufgesetzt wird, die ihn in mehr oder weniger dieselbe Position zwingt. Zur Erhöhung der Genauigkeit und der Vereinfachung der Bildaufnahme ist es allerdings wünschenswert, Positionsunterschiede mittels multimodaler Bildregistrierung des Kopf-Hals-Bereichs ausgleichen zu können. Zur Verkürzung der Planungserstellung werden rechnergestützte Bildsegmentierungsverfahren benötigt.

**Verlaufskontrolle:** Da bei dem Plattenepithelkarzinom die Rezidivquote sehr hoch ist, ist eine strikte Verlaufskontrolle nötig. Unter anderem werden dazu CT- und/oder MRT-Aufnahmen herangezogen. Dabei kommt es in erster Linie darauf an, jede Veränderung, insbesondere bei den Lymphknoten, zu erkennen. Ein direkter Vergleich von Aufnahmen aus der Zeitserie ist dabei nur schwer möglich, weil die Position des Patienten in den Aufnahmen der Zeitserie in der Regel nicht dieselbe ist und so eine nachträgliche bildbasierte Positions Korrektur benötigt wird. Um Veränderungen auch auf dem korrigierten Bild tatsächlich erkennen zu können, ist eine besondere Anforderung, dass lediglich Bildunterschiede ausgeglichen werden sollen, die aufgrund einer Positionsänderung zu Stande gekommen sind, nicht jedoch diejenigen, die von einer Anatomieveränderung stammen.

**Zusammenfassung:** Bildgebende Verfahren werden zur Diagnose, Stadienbestimmung, Therapieplanung und -verlaufskontrolle sowie der Nachverfolgung von Patienten mit Kopf-Hals-Karzinomen herangezogen. Zur Verbesserung existierender Anwendungen sowie zur Ermöglichung neuer Anwendungsgebiete werden unterschiedliche Methoden der medizinischen Bildverarbeitung benötigt. Ein Überblick ist in Tabelle 1.1 zu finden. Für Lymphknoten ergeben sich in der aktuellen klinischen Vorgehensweise folgende Nachteile:

1. Die Größenbestimmung nach den RECIST-Kriterien ist aufgrund der Betrachtung einer einzel-

nen Schicht eines 3D-Objektes ungenau und unterliegt einer hohen Inter- und Intra-Observer-Variabilität.

2. Eine manuelle Segmentierung unterliegt ebenfalls einer hohen Inter- und Intra-Observer-Variabilität. Deshalb und aufgrund des hohen Zeitaufwands zur Erstellung findet dies in der Praxis nicht statt.
3. Bei der Betrachtung von Lymphknoten über mehrere Zeitpunkte hinweg verschärft sich das Problem mangelnder Robustheit manueller Segmentierung, weil hier Volumenunterschiede einen Einfluss auf den weiteren Therapieverlauf haben können.
4. In der klinischen Praxis werden unterschiedliche Aufnahmen, seien es Aufnahmen einer Zeitreihe oder unterschiedlicher Bildmodalitäten zum selben Zeitpunkt, üblicherweise unabhängig voneinander oder bestenfalls nebeneinander betrachtet und analysiert. Eine Bildfusion findet mangels in der Praxis verfügbarer robuster deformierbarer Registrierungsverfahren nicht statt. Parameter von Lymphknoten werden also aus je einer Modalität zu je einem Zeitpunkt unabhängig bestimmt.

## 1.2. Anwendungsszenario

Um diese Nachteile überwinden zu können, wird in dieser Arbeit ein neuartiger computergestützter Ansatz zur personalisierten Vorhersage des Krankheitsverlaufs verfolgt. Für den Patienten wird ein multimodales, multitemporales parametrisches Modell erstellt (siehe Abbildung 1.2). Grundlage dafür sind Aufnahmen unterschiedlicher Bildmodalitäten, wie z.B. CT und MRT, welche komplementäre Informationen bieten. Diese werden zur Stadienbestimmung vor der Therapie akquiriert. Im Zuge der Therapieverlaufskontrolle und Nachsorge werden in Abständen von einigen Monaten weitere Aufnahmen unterschiedlicher Bildmodalitäten erstellt.

Diese Aufnahmen werden zunächst fusioniert, wodurch ein einziges Multikanal-Bild entsteht. Jeder Kanal entspricht einem anderen Zeitpunkt und/oder Bildmodalität. Voraussetzung für die Fusion ist örtliche Korrespondenz der Aufnahmen, die durch automatische Bildregistrierung hergestellt wird.

In einem weiteren Schritt werden die Zielstrukturen, also der Primärtumor und klinisch relevante Lymphknoten, definiert. Dies geschieht durch 3D-Segmentierung in mindestens einer Bildmodalität pro Zeitpunkt. Aufgrund bestehender örtlicher Korrespondenz kann diese in die anderen Kanäle propagiert werden.

Anschließend werden aus den fusionierten Bilddaten (semi-) automatisch quantitative geometrische und Textur-Merkmale aus den Zielstrukturen extrahiert [2]. Diese werden dann zur personalisierten Vorhersage des Krankheitsverlaufs mittels maschineller Klassifikationsverfahren, welche durch Patientenmerkmale mit bekanntem Krankheitsverlauf zuvor trainiert wurden, verwendet.

Zur Erhöhung der Vorhersagegenauigkeit können zu den bildbasierten Merkmalen auch klinische und genomische Merkmale in den Klassifikationsprozess einbezogen werden [3][4]. Eine verlässliche personalisierte Vorhersage des Krankheitsverlaufs ermöglicht eine Anpassung der Therapie sowie der Verlaufskontrolle. Dies kann dem Patienten unnötige operative Eingriffe mit erheblichen Nachwirkungen ersparen und zudem die Gesundheitskosten reduzieren, ohne dabei allerdings Patienten mit einem hohen Risiko eines Rezidivs unterzuversorgen.



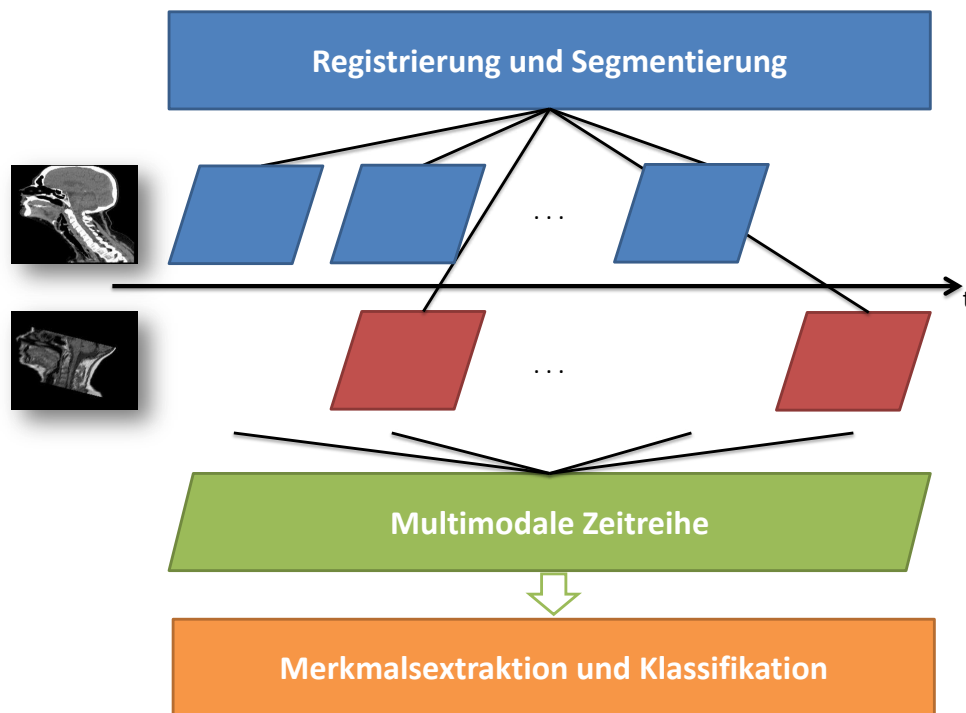


Abbildung 1.2.: Multimodales, multitemporales Modell zur Vorhersage des Krankheitsverlaufs. Aufnahmen unterschiedlicher Bildmodalitäten und Zeitpunkte werden mittels deformierbarer Bildregistrierung fusioniert. Zielorgane werden durch Bildsegmentierung definiert. Auf Grundlage dessen werden charakteristische Merkmale extrahiert, die zur personalisierten Klassifikation verwendet werden.

Um klinisch relevante bildbasierte Merkmale aus einer multimodalen Zeitreihe eines Lymphknotens extrahieren zu können, ergeben sich aus der Sicht eines Bildverarbeiters folgende zwei Herausforderungen: Die automatisierte Registrierung der Aufnahmen unterschiedlicher Bildmodalitäten und Zeitpunkten und die robuste Segmentierung von Lymphknoten aus einer Bildmodalität.

## 1.3. Zielsetzung

Ziel dieser Arbeit ist es, neue automatische Verfahren zur Registrierung und Segmentierung von Lymphknoten aus multimodalen Zeitreihen im Kopf-Hals-Bereich, die über den aktuellen Stand der Forschung hinaus gehen, zu entwickeln und zu evaluieren.

### 1.3.1. Segmentierung von Lymphknoten

Lymphknoten sollen aus medizinischen Bilddaten des Kopf-Hals-Bereichs segmentiert werden. Aufgrund der typischerweise geringeren Schichtabstände von CT- im Vergleich zu MRT-Aufnahmen

und den normierten absoluten Intensitätswerten von unterschiedlichem Gewebe soll die Segmentierung auf CT-Daten erfolgen. Eine nachfolgende Bildregistrierung kann dazu dienen, das Ergebnis auf MRT-Daten zu übertragen. Die Bestimmung des zu segmentierenden Lymphknotens soll sich auf minimale Anwenderinteraktion, also auf Vorgabe einer einzelnen 3D-Koordinate im Inneren des Lymphknotens, beschränken. Das Finden der Lymphknotengrenzen soll dann automatisch erfolgen. Bei Bedarf soll der Anwender jedoch die Möglichkeit erhalten, das Ergebnis auf einfache Art und Weise zu beeinflussen. Die Genauigkeit sollte so hoch sein, dass die Verwendung einen Mehrwert bei der Größenvermessung gegenüber RECIST darstellt. Die Laufzeit sollte einige wenige Sekunden auf aktueller Hardware nicht überschreiten. Die größten Herausforderungen bei der Segmentierung von Lymphknoten sind die hohe Varianz an Größe, Form und an umliegendem Gewebe, sowie dem oftmals geringen Kontrast zu benachbarten Strukturen und das Vorhandensein von Pathologien. Seit 1996 bis heute werden zwar von unterschiedlichen Arbeitsgruppen regelmäßig Verfahren zur Lymphknotensegmentierung vorgestellt, keines genügt jedoch allen hier definierten Anforderungen und/oder ein Teil der genannten Herausforderungen wird nicht adressiert.

### 1.3.2. Registrierung von Lymphknoten

Lymphknoten im Kopf-Hals-Bereich sollen sowohl auf verschiedenen Modalitäten als auch über Aufnahmen desselben Patienten zu unterschiedlichen Zeitpunkten hinweg registriert werden. Folgende Anforderungen ergeben sich:

- *Unterschiedliche Modalitäten:* Als Modalitäten kommen (kontrastverstärkte) CT- und MRT-Aufnahmen des Kopf-Hals-Bereichs in Frage, wobei in der klinischen Praxis übliche anatomische MR-Sequenzen (T1, T2 gewichtet, . . . ), Schichtabstände und Orientierungen der MR-Schichten (Axial, Coronal oder Sagittal) unterstützt werden sollen.
- *Unterschiedliche Bildqualitäten:* Weder typische Bildartefakte, wie z.B. Sternartefakte von Metallimplantaten in CT-Aufnahmen oder lokal unterschiedliche Signalstärke in MRT-Aufnahmen noch eine Begrenzung der Aufnahmen auf Bildausschnitte sollen ein Hindernis darstellen.
- *Unterschiedliche Patientenlage:* Bei der Positionierung des Patienten während der Aufnahme sollen keine besonderen Vorkehrungen getroffen werden müssen, was bedeutet, dass sich insbesondere die Halswirbelsäule in unterschiedlichen Stellungen in den verschiedenen Aufnahmen befinden darf.
- *Unterschiedliche Zeitpunkte:* Die Aufnahmen können zu unterschiedlichen Zeitpunkten aufgenommen sein. Insbesondere bedeutet dies, dass Änderungen in der Anatomie vorhanden sein können. Ursache dafür kann bspw. eine Gewichtsänderung, Tumorwachstum, aber auch ein chirurgischer Eingriff sein.
- *Erhaltung von Anatomieunterschieden:* Bei der Registrierung von Zeitserien ist eine weitere Anforderung, dass anatomische Unterschiede in den Lymphknotenregionen weitestgehend erhalten bleiben, so dass die Registrierung es einem Radiologen ermöglicht evtl. vorliegende Unterschiede zu erkennen.
- *Automatismus:* Die Registrierung soll vollautomatisch ablaufen. Die Segmentierung des zu registrierenden Lymphknotens kann in einer Modalität als gegeben angesehen werden.

Existierende Lösungen genügen diesen Anforderungen nicht und es werden daher neuartige Registrierungsverfahren benötigt.

## 1.4. Aufbau der Arbeit

Im folgenden Kapitel (Kapitel 2) werden zum besseren Verständnis dieser Arbeit grundlegende Eigenschaften von Lymphknoten beschrieben und gängige Bildsegmentierungs- und Registrierungsverfahren in einem globalen Kontext kategorisiert. Darauf aufbauend werden in Kapitel 3 existierende Arbeiten zum Thema Lymphknotensegmentierung und Registrierung des Kopf-Hals-Bereichs vorgestellt und theoretisch auf deren Stärken und Schwächen hin untersucht. Dabei stellt sich heraus, dass existierende Verfahren den Anforderungen nicht genügen und damit neue Verfahren entwickelt werden müssen.

In Kapitel 4 werden die eigenen Lösungsansätze zur Lymphknotensegmentierung und -registrierung motiviert und vorgestellt. Zur Lymphknotensegmentierung wird zunächst das größeninvariante Masse-Feder-Modell als Weiterentwicklung eines bestehenden modellbasierten Segmentierungsalgorithmus vorgestellt. Da dies jedoch konzeptbedingt trotz Verbesserung immer noch keine zufriedenstellenden Ergebnisse auf klinischen Daten erzielen kann, wird eine auf Radialstrahlen basierende Methode entwickelt. Besonderer Stellenwert wird dabei dem Vergleich von Kostenfunktionen und Optimierungsverfahren, sowie der Parameterwahl beigemessen, entscheidende Aspekte, welcher in der Literatur zur Lymphknotensegmentierung bisher nicht betrachtet wurden. Weitere neue Entwicklungen sind eine auf kumulierten Intensitätswerten basierende Kostenfunktion, eine auf der Bildgradientenrichtung beruhende lokale Formerhaltung und Erweiterung zur Erhöhung der Robustheit. Alternativ zur präzisen automatischen Segmentierung mittels Graph-Cut ermöglicht die Verwendung von Belief-Propagation auch eine komfortable interaktive Segmentierung.

Anschließend wird zur Registrierung von Lymphknoten ein neues Verfahren vorgestellt, welches auf einem neuartigen *artikulierten Atlas* des Skeletts basiert und sich aus insgesamt 5 Schritten zusammensetzt. In einem ersten Schritt wird der Kopf rigide registriert. Anschließend erfolgt eine Anpassung des artikulierten Atlas an die CT-Aufnahme, wodurch i) eine Segmentierung der einzelnen rigiden Strukturen (Knochen) erfolgt und ii) der artikulierte Atlas durch personalisiertes Wissen angereichert wird. Letzteres ermöglicht es, den Atlas in einem weiteren Schritt an die andere Bildmodalität anzupassen. Auf Grundlage der nun bekannten Transformation der einzelnen Knochen wird daraufhin in einem zweistufigen Prozess ein dichtes Deformationsfeld berechnet. Abschließend werden die Lymphknoten lokal rigide registriert. Dieses Verfahren ist das erste vollautomatische Verfahren seiner Art, welches sich auch auf andere Körperregionen generalisieren lässt und verfügt als eines der wenigen nicht-rigiden Registrierungsverfahren über globale Regularisierung. Die innovative Anpassungsstrategie des artikulierten Atlas führt zu einem hohen Konvergenzbereich.

Die durchgeführten Experimente bzgl. Genauigkeit, Robustheit und Laufzeit aller im Rahmen dieser Arbeit entwickelten Verfahren sowie deren qualitativen und vor allem quantitativen Ergebnisse anhand klinischer Daten werden in Kapitel 5 beschrieben.

Anschließend wird in Kapitel 6 die Anwendung der entwickelten Registrierungs- und Segmentierungsverfahren zur Merkmalsextraktion aus multimodalen Zeitreihen zur Vorhersage des Krankheitsverlaufs von Patienten mit Kopf-Hals-Karzinom beschrieben. Darüber hinaus werden weitere und denkbare Anwendungsgebiete der in dieser Arbeit entstandenen Technologien beschrieben.

## *1. Einleitung*

---

Zur Abrundung werden in Kapitel 7 die Ergebnisse der Arbeit mit Fokus auf die eigenen wissenschaftlichen Beiträge zusammengefasst und diskutiert. Abschließend wird ein kurzer Ausblick auf mögliche sinnvolle Weiterentwicklungen gegeben.

## 2. Grundlagen

### 2.1. Lymphknoten

Lymphknoten sind ein wichtiger Bestandteil des lymphatischen Systems des Menschen, welches neben der Immunabwehr auch dem Flüssigkeitstransport dient. Dort filtern sie das darin befindliche Gewebwasser (Lymphe) und befreien diese von Antigenen. In ihnen findet ein großer Teil der Immunabwehr statt.

In diesem Abschnitt wird auf die zur Registrierung und Segmentierung wichtigsten Eigenschaften von Lymphknoten, deren krankhaften Veränderungen und deren Darstellung mittels bildgebender Verfahren eingegangen. Einige Aspekte werden durch die Analyse einer repräsentativen Menge von 49 Lymphknoten aus CT-Aufnahmen untermauert. Diese Menge enthält normale, vergrößerte Lymphknoten, sowie Lymphknoten mit zentralem nekrotischem Gewebe. Einige der Aufnahmen sind kontrastmittelverstärkt. Für alle Lymphknoten liegt eine Referenz-Segmentierung vor. Eine genaue Beschreibung dieser Menge an Lymphknoten ist in Kapitel 5.1.1 zu finden.

#### 2.1.1. Eigenschaften

**Vorkommen:** Im menschlichen Körper befinden sich ca. 800 Lymphknoten, die sich mit Ausnahme des zentralen Nervensystems im gesamten Körper verteilen. Dabei gibt es Regionen mit besonderer Häufung, wie die Achselhöhle, die Bauchwand (Leiste), der Bauch, der Brustraum und der Hals. In Letzterem befinden sich um die 300 Lymphknoten [27]. Zur regionalen Einteilung der Lymphknoten im Halsbereich wird heutzutage üblicherweise das Klassifikationsschema nach Robbins et al. [28] herangezogen, welches auf der im Jahre 1932 von Rouvière [29] vorgestellten Einteilung basiert. Diese Unterteilung erfolgt dabei in den sechs in Abbildung 2.1(a) dargestellten Regionen.

**Aufbau:** Wie in Abbildung 2.1(b) zu sehen ist, gliedern sich Lymphknoten in drei Bereiche: Rinde (Kortex), Parakortikalzone und Mark [30]. Diese sind von einer Kapsel umgeben. Die Lymphe eines bestimmten Körpergebietes oder Organs um den Lymphknoten – das sog. *Einzugsgebiet* des Lymphknotens – gelangt über mehrere lymphatische *Vas afferens* in den Lymphknoten hinein. Innerhalb des Lymphknotens wird die Lymphe gefiltert und dabei von Antigenen befreit. Zusätzlich wird die Lymphe durch das Mark mit Antikörpern angereichert. Schließlich verlässt diese den Lymphknoten wieder am *Hilum* durch die lymphatische *Vas efferens*.

**Größe:** Der Durchmesser von gesunden Lymphknoten liegt normalerweise im Bereich von 2 bis 20 mm [30]. Sind sie allerdings mit der Abwehr von Krankheiten beschäftigt oder krankhaft verändert, können sie stark anschwellen. Der Durchmesser kann dann sogar 5 cm und mehr betragen. In der

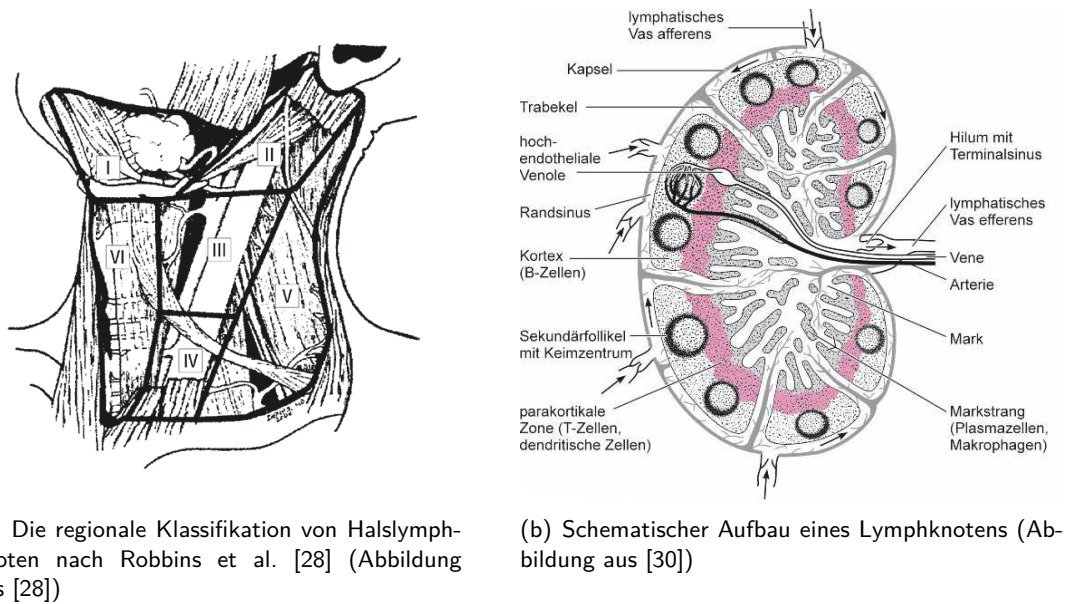


Abbildung 2.1.: Lymphknoten im Kopf-Hals-Bereich

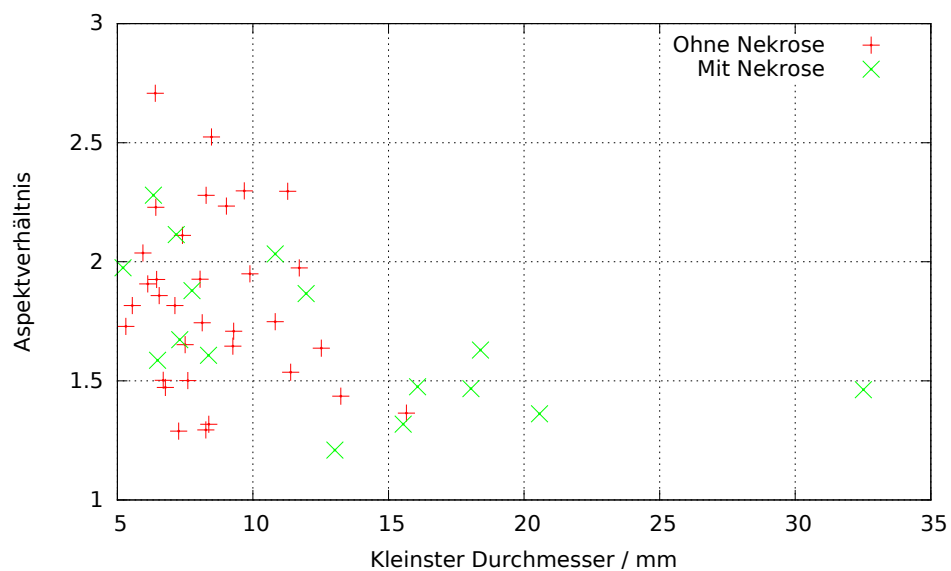


Abbildung 2.2.: Lymphknoten besitzen eine große Variabilität in Form und Größe. Je größer ein Lymphknoten ist desto runder ist er in der Regel

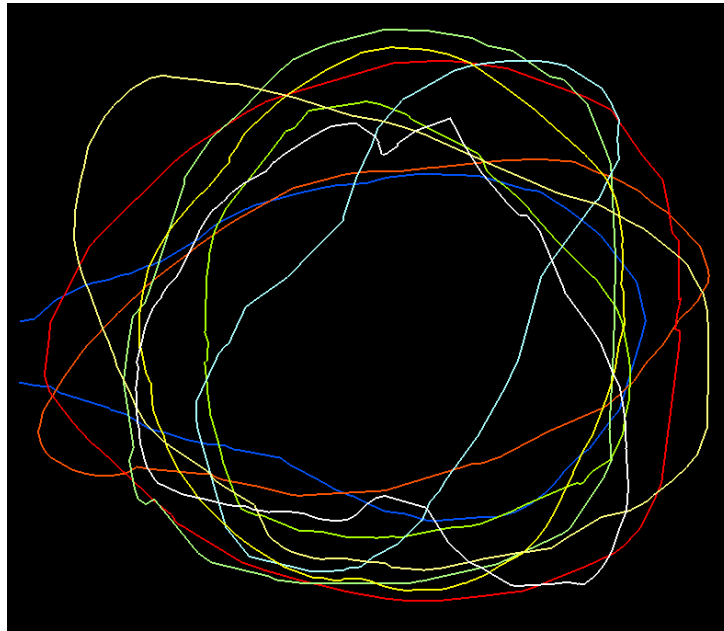


Abbildung 2.3.: Schnitt durch größennormierte Lymphknoten verdeutlicht die große Formvariabilität

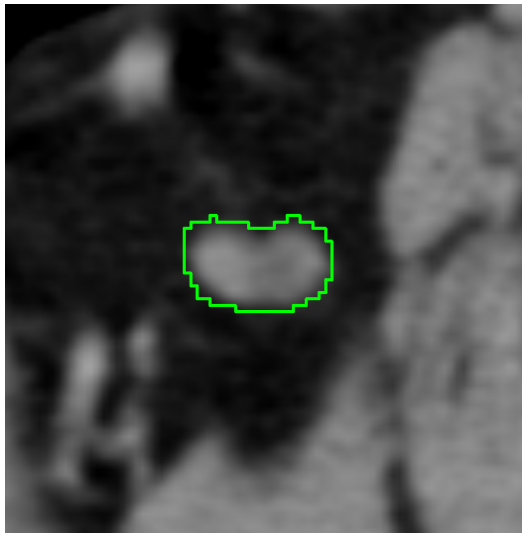
vorliegenden Menge befindet sich der große Durchmesser zwischen 9.2 und 47.6 mm (Durchschnitt  $17.1 \pm 6.9$  mm).

**Form:** Gesunde Lymphknoten sind in etwa ellipsoid bis bohnenförmig. Allerdings können einzelne Lymphknoten von dieser Form lokal stark abweichen (siehe Abbildung 2.3). Insgesamt kann die Formvariabilität von Lymphknoten daher als hoch angesehen werden. Das Verhältnis zwischen großem und kleinem Durchmesser beträgt größenordnungsmäßig 2 : 1. Wie anhand der vorliegenden Menge in Abbildung 2.2 zu sehen ist, neigen vergrößerte Lymphknoten eher dazu rund zu sein. Das Verhältnis der Durchmesser nähert sich also 1 : 1 an.

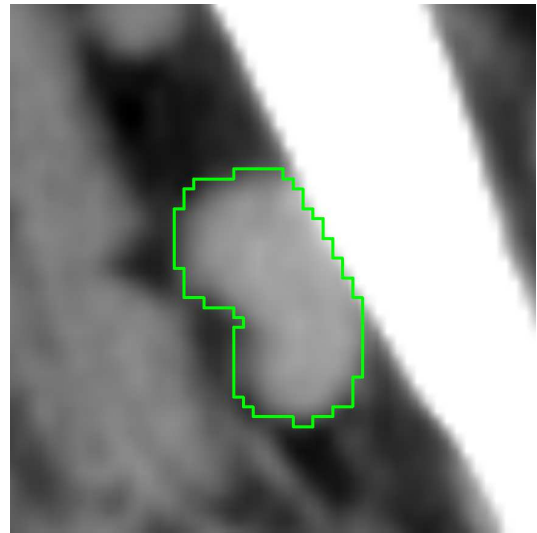
**Umliegendes Gewebe:** Lymphknoten sind oftmals von Fettgewebe umgeben. Die Fettgrenzschicht zu anderem Gewebe ist besonders im Kopf-Hals-Bereich allerdings sehr dünn, wodurch Lymphknoten de facto an sehr unterschiedliche Gewebstypen grenzen. Dieses reicht von anderem Weichteilgewebe (Gefäße und Muskeln) bis hin zu Knochen. Oftmals grenzen benachbarte Lymphknoten auch direkt aneinander an und bilden einen sogenannten Lymphknoten-Cluster. Über das umliegende Gewebe kann somit keine Annahme getroffen werden (siehe Abbildung 2.4).

### 2.1.2. Lymphadenopathie

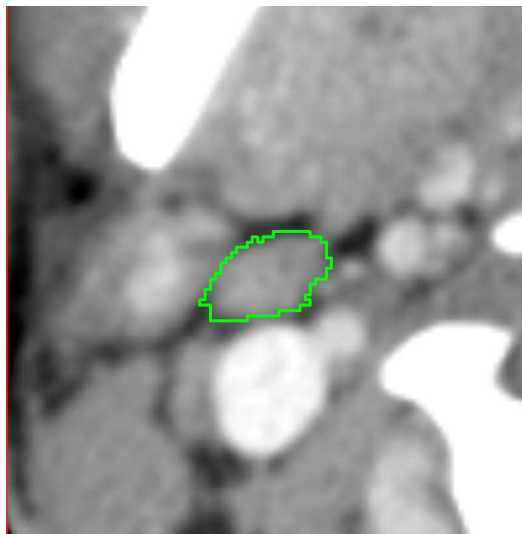
Eine krankhafte Anschwellung von Lymphknoten wird als *Lymphadenopathie* bezeichnet. Dabei wird zwischen *regionaler* und *generalisierter* Lymphadenopathie unterschieden, wobei bei Ersterer das Anschwellen der Lymphknoten auf eine bestimmte Körperregion beschränkt ist. Die Ursache für eine Lymphadenopathie kann vielfältig sein und reicht von infektiös-entzündlichen Prozessen im Einzugs-



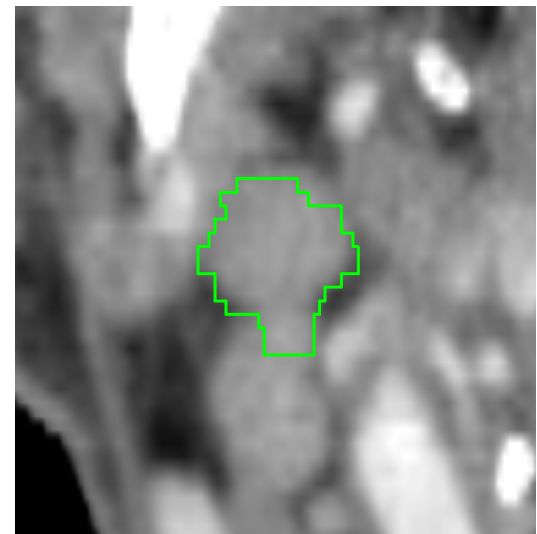
(a) Fett



(b) Knochen



(c) Gefäße und Muskeln



(d) Andere Lymphknoten

Abbildung 2.4.: CT-Aufnahmen von Lymphknoten (grün umradet) mit unterschiedlichem umliegenden Gewebe



gebiet über virale Infektionen bis hin zu bösartigen Veränderungen. Letzteres wird aufgrund des Ortes des Primärtumors weiter unterteilt. Beruht die Erkrankung auf einer bösartigen Veränderung von lymphatischen Zellen, bezeichnet man dies als *malignes Lymphom*. Sind allerdings bösartige Zellen eines Primärtumors, welcher sich im Einzugsgebiet des Lymphknotens befindet, über das lymphatische System in den Lymphknoten gelangt, spricht man von *regionalen Lymphknotenmetastasen*. Bei Krebserkrankungen im Kopf-Hals-Bereich ist das Vorhandensein von regionalen Lymphknotenmetastasen der verlässlichste prognostische Faktor für den Krankheitsverlauf [21]. Die robuste Diagnose dessen ist daher von höchster Bedeutung für die Wahl der Therapie und Verlaufskontrolle. Neben der klinischen Untersuchung und histologischen Untersuchung einer Biopsieprobe kommt dazu insbesondere auch die nichtinvasive Diagnose mittels bildgebender Verfahren zur Anwendung.

### 2.1.3. Bildgebung

Zur Untersuchung der Hals-Lymphknoten von Patienten mit Kopf-Hals-Karzinom werden unterschiedliche bildgebende Verfahren herangezogen. Neben der Computertomographie (CT) und der Magnetresonanztomographie (MRT) kommen seltener auch Ultraschall und Positronen-Emissions-Tomographie (PET) zum Einsatz. Das Ziel ist dabei das Feststellen bzw. das Ausschließen von regionalen Lymphknotenmetastasen. Obwohl, wie im vorangegangenen Abschnitt beschrieben, eine Vergrößerung von Lymphknoten unterschiedliche Ursachen haben kann, wird die Größe dennoch nach wie vor oft als Kriterium verwendet. Nach den aktuellen RECIST-Kriterien [26] gelten Lymphknoten mit einem kleinen Durchmesser von weniger als 10 mm als nicht pathologisch. Allerdings können sich Metastasen auch in kleineren Lymphknoten gebildet haben [31]. Neben der Größe wird daher einer ganzen Reihe anderer Faktoren eine Bedeutung beigemessen [31][32]. Eine ausgeprägte Bohnenform sowie die Sichtbarkeit des Hilums spricht gegen Metastasierung. Nekrotisches Gewebe innerhalb des Lymphknotens, extrakapsuläre Ausbreitung sowie ein geringer Blutfluss sind hingegen Anzeichen für regionale Lymphknotenmetastasen.

**Computertomographie:** Im Bereich der Bildgebung zur Beurteilung von Krebs im Kopf-Hals-Bereich kann das CT als das *Arbeitspferd* angesehen werden [22]. Zur Begutachtung von Lymphknoten wird dem Patienten vor der Aufnahme üblicherweise ein Kontrastmittel intravenös verabreicht, welches durchflossenes Gewebe im CT heller darstellt. Im Vergleich zu Gefäßen ist die Aufnahme von Kontrastmittel in Lymphknoten allerdings sehr gering. Zur maschinellen Verarbeitung durch Bildverarbeitungs- und Bildanalyse-Algorithmen eignet sich das CT besonders dadurch, dass der Schichtabstand mit einer Größenordnung von 1 mm relativ gering ist und die Signalstärke durch die Hounsfield-Skala (Luft  $\hat{=}$  -1000 HU, Wasser  $\hat{=}$  0 HU) normiert ist.

Die Grauwertverteilung der vorliegenden Menge ist in Abbildung 2.5 dargestellt. 95% aller zentral gelegenen Voxel – hier alle enthaltenen Voxel mit einem Mindestabstand zum Rand von 1 mm – aller nicht-nekrotischen Lymphknoten befinden sich zwischen 43 und 119 HU. Der Durchschnitt liegt bei  $80 \pm 20$  HU. Wird der äußere Bereich mit einbezogen, sinkt der Durchschnitt durch den dort zum Tragen kommenden *Partialvolumeneffekt* auf  $63 \pm 37$  HU. Lymphknoten mit zentralem nekrotischen Gewebe (siehe Abbildung 2.6(b)) haben einen deutlich dunkleren inneren Bereich.

Der Kontrast von Lymphknoten zum umliegenden Gewebe ist an vielen Stellen (benachbarte Lymphknoten, Muskeln, Gefäßen) sehr gering. Lediglich an Übergängen zu Fettgewebe ist ein größerer, beim Übergang zu Knochen und Luft sogar ein sehr großer Kontrast gegeben. Abbildung 2.6 zeigt

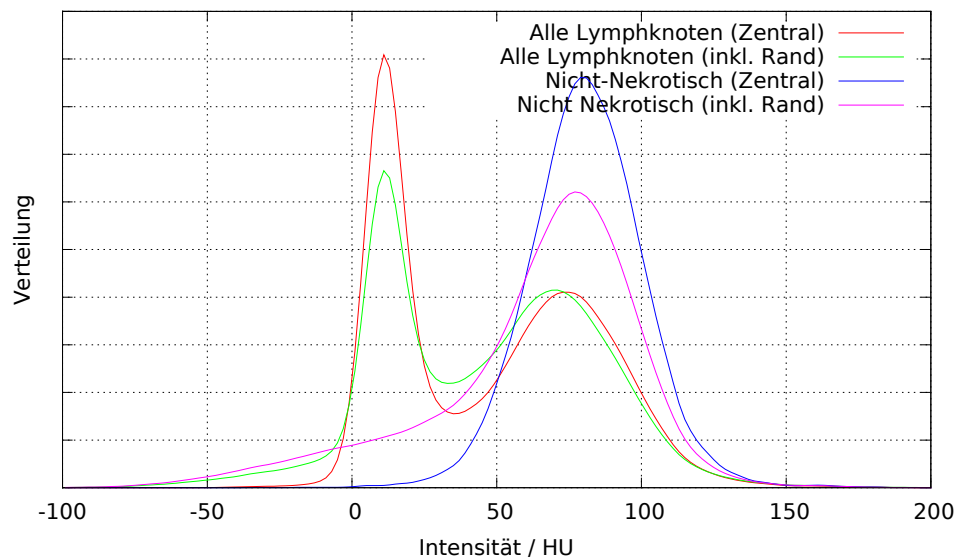


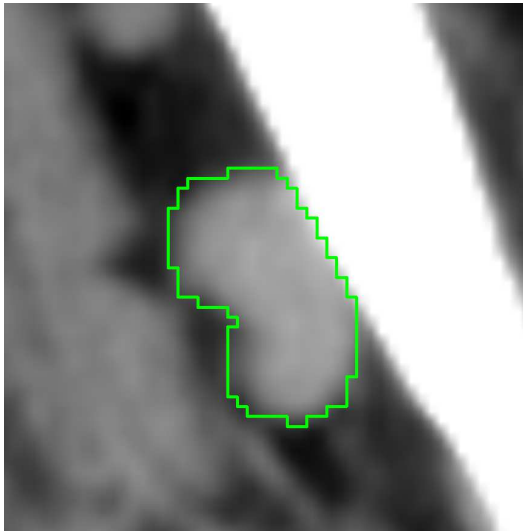
Abbildung 2.5.: Grauwertverteilung von Lymphknoten in kontrastverstärkten CT-Aufnahmen, wie sie im klinischen Alltag zum Einsatz kommen. Wird der Rand (1 mm) mit einbezogen, verschiebt sich die Verteilung aufgrund des Partialvolumeneffektes etwas zu Gunsten von geringen Intensitäten.

CT-Aufnahmen von Lymphknoten mit unterschiedlichen Eigenschaften.

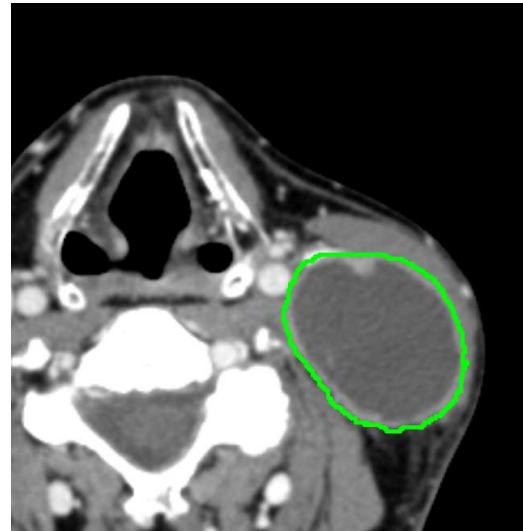
Neben der Bestimmung der Größe eignet sich die Computertomographie insbesondere zur Erkennung von nekrotischem Gewebe innerhalb des Lymphknotens [22].

**Magnetresonanztomographie:** Der große Vorteil – sowohl zur Diagnose als auch zur Bildverarbeitung – von MRT ist der hohe Weichteilkontrast (siehe Abbildung 2.7). Dem gegenüber stehen allerdings eine lange Aufnahmezeit und die damit verbundene Anfälligkeit für Bewegungsartefakte, ein großer Schichtabstand ( $> 3 \text{ mm}$ ) sowie nicht normierte Bildintensitäten. Diese werden maßgeblich von dem verwendeten MRT-Protokoll bestimmt (es steht eine sehr große Zahl an unterschiedlichen nicht standardisierten Sequenzen zur Auswahl), aber auch andere Faktoren, wie z.B. die Masse des Patienten und die lokale Magnetfeldstärke, haben einen Einfluss. Dies ist insbesondere für eine maschinelle Verarbeitung von großem Nachteil.

Neben der üblichen T1- und T2-gewichteten Aufnahmen mit Fettsättigung kommt auch ultrakleines superparamagnetisches Eisenoxid (USPI) als Kontrastmittel zum Einsatz [32], welches in T2-gewichteten Aufnahmen zu einer Reduzierung des MRT-Signals innerhalb von gesunden und entzündeten Lymphknoten führt. Metastasierte Lymphknoten zeigen diesen Signalverlust nicht, wodurch diese von gutartigen Lymphknoten unterschieden werden können. Darüber hinaus kann zur Erkennung von Lymphknotenmetastasen auch diffusionsgewichtete Magnetresonanztomografie (DWI) verwendet werden, welche die Diffusionsbewegung von Wassermolekülen in Körpergewebe misst. Eine Kenngröße des DWI-Bildes – der scheinbare Diffusionskoeffizient (ADC) – ist in metastasierten Lymphknoten höher als in gutartigen Lymphknoten sowie Lymphomen.



(a) Typischer gutartiger bohnenförmiger Lymphknoten



(b) Vergrößerter Lymphknoten mit zentralem nekrotischen Gewebe

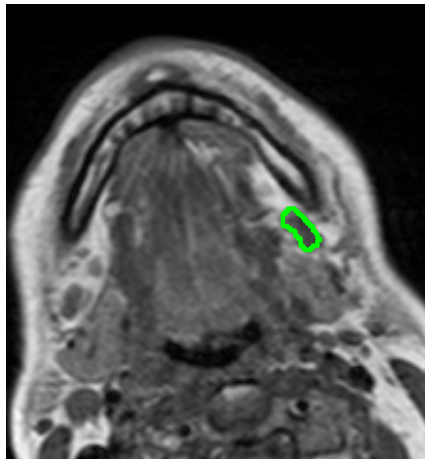


(c) Stark verrauschte Aufnahme

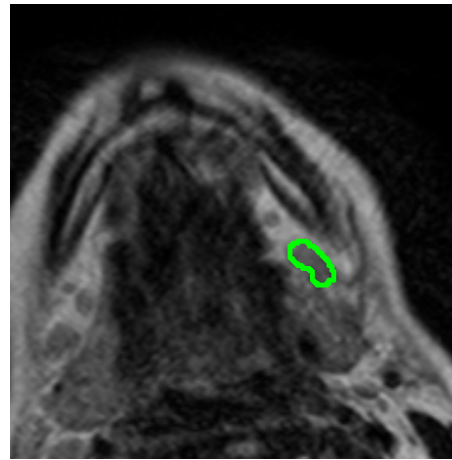


(d) Geringer Kontrast zu umliegendem Gewebe

Abbildung 2.6.: CT-Aufnahmen von Lymphknoten (grün umradet) mit unterschiedlichen Eigenschaften



(a) T1 gewichtete MRT-Aufnahme



(b) T2 gewichtete MRT-Aufnahme

Abbildung 2.7.: MRT-Aufnahmen von Lymphknoten (grün umradet)

**Ultraschall:** Lymphknoten, die nahe an der Hautoberfläche liegen, können mittels Ultraschall begutachtet werden. Die Größe und Form können bestimmt werden. Zur Begutachtung des Hilums ist Ultraschall die Methode der Wahl [31]. Dopplersonografie kann zur Flussmessung verwendet werden.

**Positronen-Emissions-Tomographie:** PET besitzt eine hohe Sensitivität zur Detektion von Lymphknotenmetastasen, allerdings geht diese mit einer geringen Spezifität einher [21], was zu einer Falschklassifikation tatsächlich nicht befallener Lymphknoten führt. Auch ist es mit PET auflösungsbedingt nicht möglich, kleinste Metastasen zu erkennen. PET spielt daher nur eine untergeordnete Rolle.

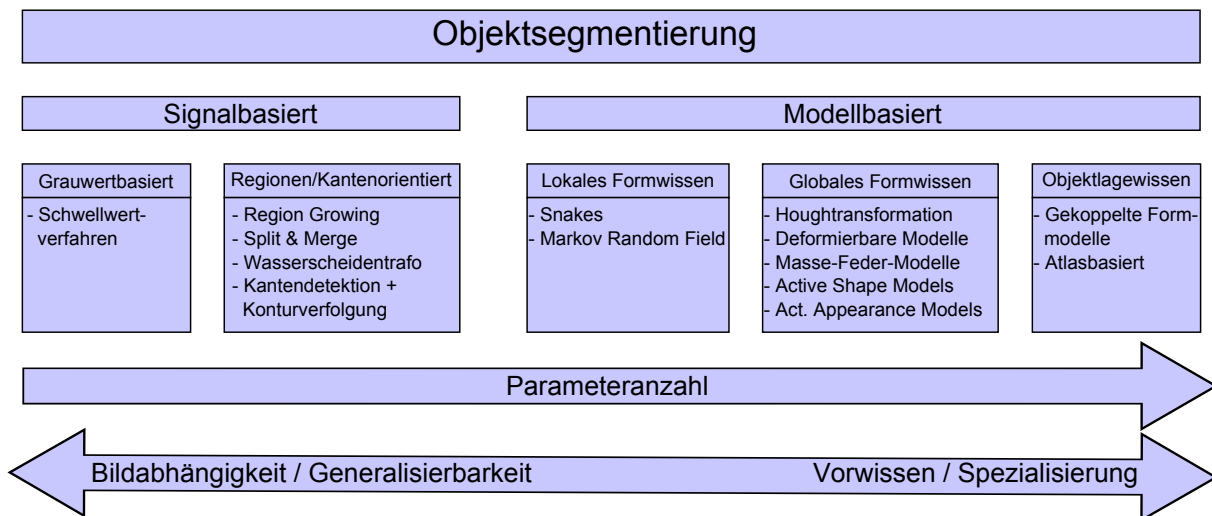


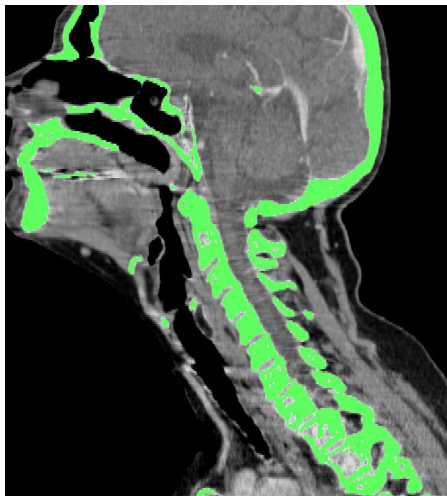
Abbildung 2.8.: Klassifikationsschema von Segmentierungsverfahren anhand der Menge des verwendeten Vorwissens

## 2.2. Segmentierung und Registrierung in der medizinischen Bildverarbeitung

Trotz jahrzehntelanger Forschung sind Bildsegmentierung und -registrierung nach wie vor die aktivsten Forschungsgebiete im Bereich der medizinischen Bildverarbeitung. Der Literaturstudie [5] zufolge, die mein Kollege Marius Erdt zusammen mit mir durchführte, stellen 45% aller in den wichtigsten Fachzeitschriften und Konferenzen vorgestellten Arbeiten der letzten zwei Jahre einen neuen Segmentierungs- und Registrierungsansatz vor, der meistens auf einen ganz speziellen Anwendungsfall abzielt. Zum besseren Verständnis dieser Arbeit werden in diesem Abschnitt kurz gängige Bildsegmentierungs- und Registrierungsverfahren kategorisiert. Insbesondere werden dabei die für diese Arbeit so wichtigen Gemeinsamkeiten dieser scheinbar getrennten Forschungsgebiete herausgestellt. Es findet allerdings keine lehrbuchartige Einführung statt. Dazu sei der Leser z.B. an das Standardwerk von Handels [33] verwiesen.

### 2.2.1. Segmentierung

Unter Bildsegmentierung versteht man die Aufteilung eines Bildes in zusammenhängende Bereiche. Sie wird in der Bildverarbeitung üblicherweise zur Lokalisierung von Objekten verwendet. Die im Jahr 2000 von Duncan et al. [34] getroffene Feststellung, dass kein einzelner Algorithmus in der Lage ist, eine Vielzahl von relevanten Strukturen in medizinischen Bildern zu segmentieren, besitzt nach wie vor Gültigkeit. Die großen Herausforderungen sind geringer Kontrast zwischen den einzelnen Strukturen, große Varianz bezüglich der Form und Erscheinung von Strukturen derselben Klasse sowie Bildstörungen, wie z.B. Rauschen oder Aufnahmeartefakte. Dies hat zur Folge, dass selbst trainierte Personen, wie z.B. Radiologen mit jahrelanger Erfahrung in vielen Fällen eine hohe Intra- und Inter-Observer-Variabilität aufweisen. Zur Segmentierung eines Objektes aus einem medizinischen Bild können vier verschiedene Merkmalstypen herangezogen werden: Diskontinuität am Übergang zu



(a) Segmentierung der Knochen mittels Schwellwertverfahrens. Das Ergebnis ist an einigen Stellen lückig.



(b) Segmentierung eines Lymphknoten mittels Region-Growing. Die Segmentierung läuft in benachbarte Strukturen aus.

Abbildung 2.9.: Signalbasierte Segmentierungsverfahren angewendet auf einen CT-Datensatz

anderen Objekten, Erscheinung, Form und Lage. Nicht alle Verfahren verwenden alle vier Charakteristiken. Je mehr Wissen ein Verfahren über Erscheinung, Form und Lage im Segmentierungsprozess einbezieht, desto geringer ist die Empfindlichkeit gegenüber geringem Bildkontrast und Bildstörungen. Gleichzeitig ist es dabei jedoch auch spezieller auf eine bestimmte Anwendung zugeschnitten. Abbildung 2.8 zeigt ein Schema, welches gängige Segmentierungsverfahren in fünf Hauptkategorien innerhalb des Kontinuums bzgl. der Menge des verwendeten Vorwissens und damit der Spezialisierung einordnet. Im Folgenden wird auf diese Hauptkategorien kurz eingegangen.

**Grauwertbasierte Methoden:** Grauwertbasierte Methoden entscheiden die Objektzugehörigkeit für jeden einzelnen Pixel/Voxel unabhängig, einzig und allein aufgrund der Bildintensität. Nur Pixel/Voxel mit Grauwerten innerhalb eines bestimmten Intervalls oder über/unter einem bestimmten Schwellwert werden zu dem zu segmentierenden Objekt gezählt. Das zugrundeliegende Bild muss dabei nicht zwangsläufig das Originalbild, sondern kann auch eine vorverarbeitete Version, sein. Beispielsweise kann durch lineare Filterung Rauschen unterdrückt oder gewisse Strukturen hervorgehoben werden. Konzeptbedingt funktionieren grauwertbasierte Verfahren allerdings nur bei Strukturen mit hohem Bildkontrast. Das prominenteste Anwendungsgebiet ist die Segmentierung von Knochen in CT-Aufnahmen (siehe Abbildung 2.9(a)). Das einzige Vorwissen, das dabei zur Anwendung kommen kann, ist, in welchem Grauwertbereich sich die zu segmentierende Struktur befindet. Dieser kann entweder von einem Experten auf dem Anwendungsgebiet vorgegeben werden oder durch maschinelle Lernverfahren anhand von manuell segmentierten Trainingsdaten gelernt werden. Alternativ kann die Grauwertschwelle auch automatisch aufgrund der Grauwertverteilung im Bild bestimmt werden. Dies geschieht oftmals durch das Otsu-Verfahren [35], welches den Schwellwert so setzt, dass die Grauwertvarianz innerhalb der Objektklassen möglichst gering und zwischen den Objektklassen möglichst hoch ist.

**Regionen- und kantenorientierte Verfahren:** Regionen- und kantenorientierte Verfahren sind wie die grauwertbasierten Verfahren ebenfalls signalbasiert. Allerdings werden Nachbarschaftsbeziehungen in den Segmentierungsprozess einbezogen.

Beim *Region-Growing* wird ausgehend von einem Saatpunkt eine Region solange erweitert, bis ein anwendungsspezifisches Homogenitätskriterium nicht mehr erfüllt ist. Durch diese Vorgehensweise ist garantiert, dass die entstandene Region zusammenhängend ist. Da allerdings keinerlei Forminformation einbezogen wird, passiert es oft, dass die Region an ein oder mehreren Stellen aus dem zu segmentierenden Objekt ausläuft (*Boundary-Leaking*, siehe Abb. 2.9(b)). Die Wahl des Saatpunktes erfolgt meist durch den Anwender. Aufgrund der geringen Komplexität und der Generalisierbarkeit des Ansatzes hat sich Region-Growing zu einem Standardtool in kommerzieller Bildverarbeitungssoftware entwickelt. Weitere oft angewendete regionenbasierte Verfahren sind der *Split-and-Merge-Algorithmus* [36] und die *Wasserscheidentransformation* [37].

Kantenorientierte Verfahren basieren auf der Detektion von Diskontinuitäten im Bild am Übergang zwischen zwei Objekten. Dies geschieht üblicherweise durch Gradientenfilter oder darauf aufbauenden Methoden, wie z.B. dem *Canny-Edge-Detector*[38]. Da die Konturen im Allgemeinen nicht geschlossen sind, sind oftmals weitere Verarbeitungsschritte nötig, bevor ein Konturverfolgungsalgorithmus eingesetzt werden kann.

**Verfahren mit lokalem Formwissen:** Die erste Verfahrenskategorie, die sich nicht ausschließlich auf die Diskontinuität zwischen Objekten und/oder dem Erscheinungsbild des zu segmentierenden Objektes bezieht, ist die Segmentierung mit lokalem Formwissen. Hierbei wird das Segmentierungsproblem meist in ein Energieminimierungsproblem überführt, das diejenige Segmentierung  $S^*$  findet, welche ein Energiefunktional der Form

$$E(S) = E_{\text{extern}}(B, S) + \lambda \cdot E_{\text{intern}}(S) \quad (2.1)$$

minimiert.  $E_{\text{extern}}(B, S)$  gibt an, wie gut die Segmentierung  $S$  zu dem Bild  $B$  passt, also bspw. wie stark die Bildgradienten am Rand von  $S$  sind, wohingegen  $E_{\text{intern}}(S)$  unwahrscheinliche, also bspw. sprunghafte Formen bestraft.  $\lambda$  regelt den Einfluss der einzelnen Terme. Die Segmentierung  $S$  kann auf unterschiedliche Weise parametrisiert werden. Bei den 1988 von Kass et al. vorgestellten *Snakes* [39] zur 2D-Segmentierung von glatten Objekten wird eine (geschlossene) Kontur  $x(\nu)$  durch einen skalaren Wert  $\nu \in [0, 1]$  repräsentiert. In 3D werden diese oft als *Aktive Konturen* bezeichnet und können z.B. durch Polygonnetze beschrieben werden. In [40] wird implizite Modellierung durch Niveaumengen (*Level-Sets*) vorgestellt. Ein (lokales) Energieminimum wird meist durch ein lokales Optimierungsverfahren gefunden. Der Segmentierungserfolg hängt daher zu einem großen Teil von der Initialisierung ab. Insbesondere wenn das Segmentierungsproblem als *Markov-Random-Field* (MRF) modelliert werden kann, können aber auch globale Optimierungsverfahren wie *Graph-Cut* [41] oder *Belief-Propagation* [42][43] zum Einsatz kommen.

Das Einbeziehen von lokalem Formwissen hat gegenüber rein signalbasierten Methoden den Vorteil, dass Objektgrenzen mit lokal sehr geringem Kontrast aufgrund von kontrastreichen Objektgrenzen in der Umgebung dennoch gefunden werden können. Da die möglichen Formen gegenüber signalbasierten Methoden eingeschränkt sind, sind Verfahren mit lokalem Formwissen deutlich spezieller an ein bestimmtes Segmentierungsproblem angepasst.

**Verfahren mit globalem Formwissen:** Bei Verfahren mit globalem Formwissen wird die interne Energie  $E_{\text{intern}}$  aus Gleichung 2.1 dahingehend erweitert, dass der Unterschied zu einer Referenzform oder einer Menge von Referenzformen einfließt. Es können also nur noch Objekte, die so ähnlich wie die Referenzform(en) aussehen, segmentiert werden. Dabei unterscheidet man zwischen *Deformierbaren Modellen* [44], welche den Unterschied zu einer einzigen Referenzform einbeziehen, *Active Shape Models*(ASMs) [45], welche einen gelernten Formenraum betrachten und *Active Appearance Models*(AAMs) [46], welche bei der Berechnung der externen Energie  $E_{\text{extern}}$  zusätzlich zur Betrachtung der Kontur auch Wissen über das Aussehen innerhalb des Modells miteinbeziehen.

Bei ASMs wird eine einzelne Form  $\mathbf{x}$  oftmals mittels Hauptkomponentenanalyse (Principal Component Analysis, PCA) mittels eines niedrigdimensionalen Vektors  $\mathbf{b}$  durch

$$\mathbf{x} = \bar{\mathbf{x}} + A \cdot \mathbf{b} + \mathbf{r} \quad (2.2)$$

beschrieben, wobei  $\bar{\mathbf{x}}$  dem Durchschnitt aller  $n$  Trainingsformen  $\mathbf{x}_j$  entspricht. Die Matrix  $A$  besteht aus den Eigenvektoren der Kovarianzmatrix  $C = \frac{1}{n-1} \bar{X} \bar{X}^T$ ,  $\bar{X} = \mathbf{x}_0 - \bar{\mathbf{x}}, \dots, \mathbf{x}_{n-1} - \bar{\mathbf{x}}$ . Der Vektor  $\mathbf{r}$  ist das Residuum. Je näher eine beliebige Form  $\mathbf{x}$  an dem durch die Trainingsformen  $\mathbf{x}_j$  beschriebenen Formenraum liegt, desto geringer sind  $|\mathbf{b}|$  und  $|\mathbf{r}|$ .

Weitere Verfahren, die zu dieser Kategorie gezählt werden können, sind *stabile Masse-Feder-Modelle* [47] und die bereits 1962 patentierte Hough-Transformation [48].

Verfahren mit globalem Formwissen kommen immer dann zum Einsatz, wenn sich das zu segmentierende Objekt nicht klar von dem Hintergrund abgrenzt, dafür aber viel über die Form bekannt ist. Sie sind also sehr speziell auf einen bestimmten Anwendungsfall, wie z.B. die Segmentierung eines einzelnen Organs [6], zugeschnitten.

**Lageorientierte Verfahren:** In einigen Fällen reicht die Einbringung von globalem Formwissen immer noch nicht aus, um ein bestimmtes Objekt zu segmentieren. Ein Beispiel dafür ist die Segmentierung eines bestimmten Wirbels. Da die Formen benachbarter Wirbel üblicherweise recht ähnlich sind, kann mit Formwissen allein nicht garantiert werden, dass auch tatsächlich der gewünschte und nicht ein benachbarter Wirbel segmentiert wird. Das Einbeziehen von Vorwissen über den Ort des Objektes bzw. die relative Lage zu anderen Objekten kann hierbei behilflich sein. *Gekoppelte Modelle* [49] segmentieren mehrere Strukturen gleichzeitig unter der Einbeziehung ihrer relativen Lage. Auf die Spitze wird dieses Prinzip bei der *atlasbasierten Segmentierung* [50] gebracht (siehe Abbildung 2.10). Dort werden zunächst in einem Modellbild – dem Atlas – alle relevanten Objekte meist manuell oder interaktiv segmentiert. Anschließend wird mittels eines bildbasierten Registrierungsverfahrens die örtliche Korrespondenz des gesamten Atlas zu dem Referenzbild gefunden, für welches die Segmentierung erfolgen soll. Der eigentliche Segmentierungsschritt besteht nun lediglich darin, die Segmentierung aus dem Atlas zu übertragen. Die Güte der Segmentierung hängt von der Auswahl des Atlas und dem verwendeten Registrierungsalgorithmus ab. Die relativen Lagen der Objekte sind implizit durch den Atlas kodiert.

### 2.2.2. Registrierung

Bildregistrierung bezeichnet die Bestimmung des örtlichen Zusammenhangs zwischen zwei Aufnahmen derselben Szene oder ähnlicher Szenen [51]. Trotz der enormen Fortschritte in den vergan-



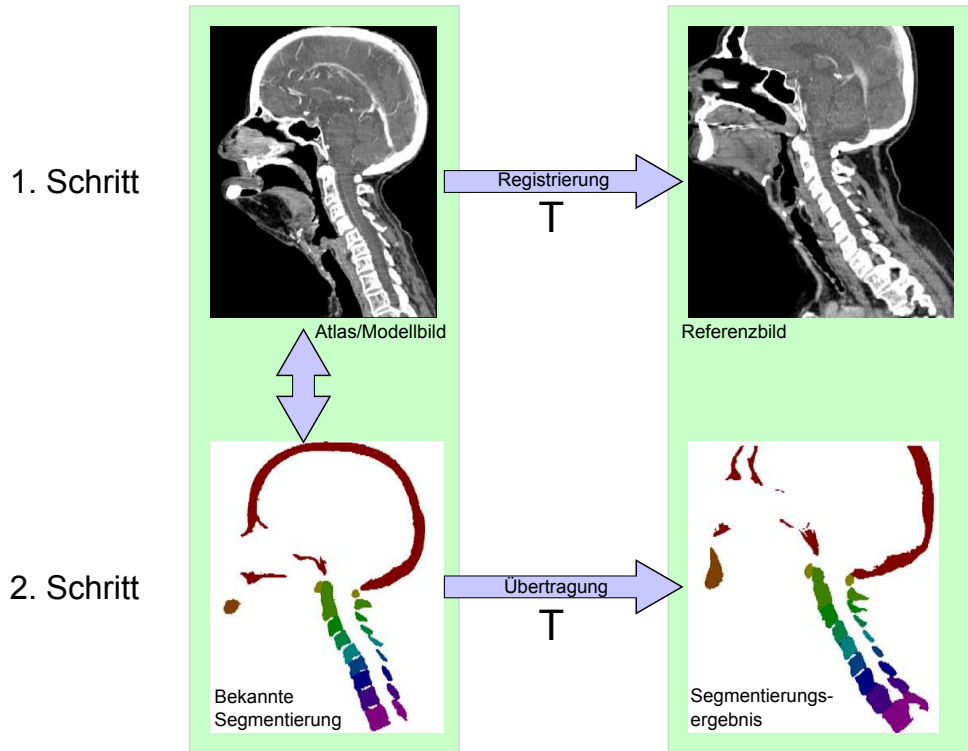


Abbildung 2.10.: Bei der atlasbasierten Segmentierung wird das Referenzbild mit einem zuvor segmentierten Modellbild (Atlas) registriert. Ist die örtliche Korrespondenz bekannt, kann die Segmentierung einfach übertragen werden.

genen Jahren hat das bereits im Jahre 1998 aufgestellte Klassifikationsschema für Bildregistrierungsverfahren für medizinische Bilder nach Maintz et al. [52] weitestgehend Gültigkeit behalten. Demnach können Registrierungsverfahren nach *Dimensionalität*, *Registrierungsgrundlage*, *Art der Transformation*, *Transformationsdomäne*, *Interaktion*, *Optimierungsverfahren*, *beteiligte Modalitäten*, *Subjekt* und *Objekt* eingeteilt werden. Ein Registrierungsproblem kann immer auch als ein Optimierungsproblem aufgefasst werden, bei dem die optimalen Parameter  $\mathbf{p}^* \in \mathbb{R}^n$  einer Transformation  $T : \mathbb{R}^3 \times \mathbb{R}^n \mapsto \mathbb{R}^3$  gefunden werden sollen, die ein Fehlermaß (Unähnlichkeitsmaß)  $F : (\mathbb{R}^3 \mapsto \mathbb{R}) \times (\mathbb{R}^3 \mapsto \mathbb{R}) \mapsto \mathbb{R}_0^+$  zwischen einem *Referenzbild*  $R$  und einem durch  $\mathbf{p}$  parametrisiertem transformierten *Modellbild*  $M_T$  der Form

$$E(\mathbf{p}) = F(R, M_T(\mathbf{p})) \quad (2.3)$$

minimiert. Maßgeblich für die zur Verfügung stehenden Möglichkeiten, aber auch der damit einhergehenden Komplexität, ist die Art der Transformation. Neben der Erhöhung der Komplexität haben flexiblere Transformationen den weiteren entscheidenden Nachteil, dass die Anforderung an das Fehlermaß steigt, weil lokale Ungenauigkeiten sich stärker bemerkbar machen. Dies kommt besonders bei multimodalen Registrierungsproblemen zum Tragen, weil in diesem Fall die beteiligten Bilder komplementäre Informationen zeigen, die aber dennoch örtlich korrekt aufeinander abgebildet werden sollen. Aber auch Bilder mit Aufnahmeartefakten, wie z.B. CT-Aufnahmen mit Metallartefakten oder MRT-Aufnahmen mit lokal geringer Signalstärke, sind davon betroffen. Die Transformation soll-

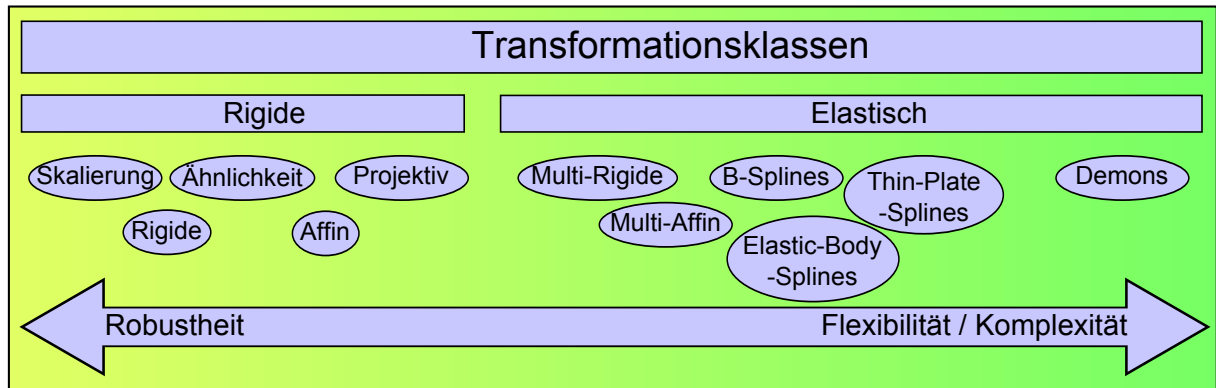


Abbildung 2.11.: Die Flexibilität von Transformationsklassen, welche in der Bildregistrierung verwendet werden, bilden ein Kontinuum.

te daher immer so unflexibel wie möglich gewählt werden, aber dennoch den Anforderungen bzgl. der zu erwartenden Deformation zwischen den Bildern und der geforderten Genauigkeit entsprechen.

### 2.2.2.1. Transformationsklassen

Als Transformation stehen in der Literatur eine ganze Reihe von Möglichkeiten zur Auswahl, welche üblicherweise in zwei Kategorien eingeteilt werden: *rigide* (globale) und *nicht-rigide* (elastische) Transformationen [51]. Diese Unterteilung ist allerdings sehr grob und wird daher im Folgenden etwas genauer dargelegt. Die Art der Unterteilung erfolgt nach dem Grad der Flexibilität (siehe Abbildung 2.11), wobei an dem einen Ende des Kontinuums eine einfache isotrope Skalierung (ein Freiheitsgrad) und am anderen Ende ein beliebiges Deformationsfeld steht, bei dem jedem der  $N$  Voxel im Referenzbild ein eigener 3-Dimensionaler Deformationsvektor zugeordnet wird ( $3N$  Freiheitsgrade).

**Rigide Transformationen:** Üblicherweise werden zu den globalen Transformationen lediglich lineare Abbildungen der Art

$$\mathbf{x}' = T(\mathbf{x}, \mathbf{p}) = A(\mathbf{p}) \cdot \mathbf{x} \quad (2.4)$$

gezählt, wobei  $A \in \mathbb{R}^{4 \times 4}$  von dem Parametervektor  $\mathbf{p}$  abhängt und sowohl  $\mathbf{x}$  als auch  $\mathbf{x}'$  in homogenen Koordinaten gegeben sind. Man beachte, dass  $A$  nicht vom Ort  $\mathbf{x}$  abhängt und so eine globale Transformation gegeben ist, also eine Transformation, die an allen Orten dieselbe ist. Die Parametrisierung von  $A$  bestimmt die Art der Abbildung, angefangen bei der *Isotropen Skalierung* mit einem Freiheitsgrad (DOF), über die *Rigide Transformation* (3 DOF Rotation + 3 DOF Translation) bis hin zur *Projektiven Abbildung* mit 15 DOF.

Nicht nur bei starren Objekten, wie z.B. dem Schädel, ist es sinnvoll, eine rigide Transformation zu verwenden. Falls nur ein kleiner Bereich von Interesse ist, kann dieser oftmals näherungsweise als starr angenommen werden. Bei der Optimierung darf das Fehlermaß dann nur in dieser zuvor bestimmten Region (ROI) berechnet werden.

Neben der geringen Komplexität und der hohen Robustheit haben rigide Verfahren den großen Vorteil, dass es nicht zu ungewünschten geometrischen Verzerrungen kommt und insbesondere Größen

erhalten bleiben. Sollen beispielsweise Anatomieänderungen in zwei Aufnahmen einer Zeitserie zur Bestimmung des Therapieerfolgs bestimmt werden, können nicht-rigide Verfahren i.d.R. nicht eingesetzt werden, da diese neben den Positionsänderungen auch etwaige Anatomieänderungen ausgleichen und so dem eigentlichen Ziel – der Ermöglichung der Unterschiedsdetektion – entgegenwirken.

**Multi-Rigide Transformationen:** Die Klasse mit der nächsthöheren Komplexität sind *multi-rigide* bzw. *multi-lineare* Verfahren. Diese Klasse von Verfahren ist die erste, die als elastisch bezeichnet werden kann. Die Grundidee basiert darauf, dass  $n$  einzelne Regionen  $R_i, i \in [1 \dots n]$  des Bildes als für sich rigide betrachtet werden. Die Bestimmung dieser rigiden Einzelteile kann dabei entweder durch gleichmäßiges Aufteilen des Bildes (z.B. [7]) oder durch die Einbringung von anatomischem Vorwissen erfolgen. Anstatt ganzer Regionen können allerdings auch einzelne Punkte eine lineare Transformation zugeordnet werden, aus denen dann ein invertierbares Deformationsfeld bestimmt wird [53].

Als anatomisches Vorwissen wird üblicherweise verwendet, dass einzelne Knochen starre Körper sind. Analog zur Physiologie des Menschen *trägt* das menschliche Skelett das umliegende Weichteilgewebe. Diese Betrachtungsweise wurde als erstes im Jahre 1996 von Little et al. [54] vorgeschlagen und seitdem immer wieder aufgegriffen. Der Registrierungsprozess besteht aus drei Schritten:

1. Segmentierung der einzelnen rigiden Strukturen,
2. Registrierung der rigiden Strukturen,
3. Bestimmung der Deformation in den restlichen Strukturen.

Neben der bereits erwähnten Analogie zur menschlichen Physiologie haben multi-rigide Transformationen den Vorteil, dass die Anzahl der Freiheitsgrade mit  $6 \cdot n$  für ein elastisches Verfahren sehr gering ist. Aufgrund dieser beiden Eigenschaften kann ein großer Konvergenzbereich (eine große *Capture Range*) erzielt werden. Die Registrierung in den starren Teilen ist sehr genau, da dort keine Verzerrungen zulässig sind. Allerdings kann es im umliegenden Gewebe zu Registrierungsungenauigkeiten kommen, da die Deformation dort üblicherweise nur durch Interpolation der rigiden Transformationen entsteht. Dies kann bspw. umgangen werden, wenn im Nachgang ein weiteres flexibleres Transformationsmodell zum Einsatz kommt.

**Verbindung von Kontrollpunkten mit Splines:** Das wohl am häufigsten zur elastischen Registrierung verwendete Transformationsmodell basiert auf der Verwendung von *Kontrollpunkten*, die zur Parametrisierung von *Splines* verwendet werden. Eine Möglichkeit zur Bestimmung des Deformationsfeldes  $\mathbf{d} : \mathbb{R}^3 \mapsto \mathbb{R}^3$  besteht darin, zunächst in dem Referenzbild  $R$  und dem Modellbild  $M$   $n$  korrespondierende Paare von *Interest-Punkten*  $(\mathbf{r}_i, \mathbf{m}_i) \in \mathbb{R}^3 \times \mathbb{R}^3, i \in [1 \dots n]$  zu detektieren, welche die Stützstellen  $\mathbf{d}_i := \mathbf{m}_i - \mathbf{r}_i$  des Deformationsfeldes  $d$  definieren ( $\mathbf{d}(\mathbf{r}_i) = \mathbf{d}_i$ ). Anschließend wird das restliche Deformationsfeld durch Interpolation bestimmt. Neben *Elastic-Body-Splines* (EBS) [55], werden dazu häufig *Thin-Plate-Splines* (TPS)[56] verwendet, welche das physikalische Verhalten einer unendlich großen, dünnen Metallplatte unter der Last von  $n$  Massen modellieren. Das Deformationsfeld ist durch

$$\mathbf{d}(\mathbf{x}) = a + b \cdot x_1 + c \cdot x_2 + d \cdot x_3 + \sum_{i=1}^n f_i \cdot (||\mathbf{x} - \mathbf{r}_i||^2 + \eta^2) \cdot \ln (||\mathbf{x} - \mathbf{r}_i||^2 + \eta^2) \quad (2.5)$$

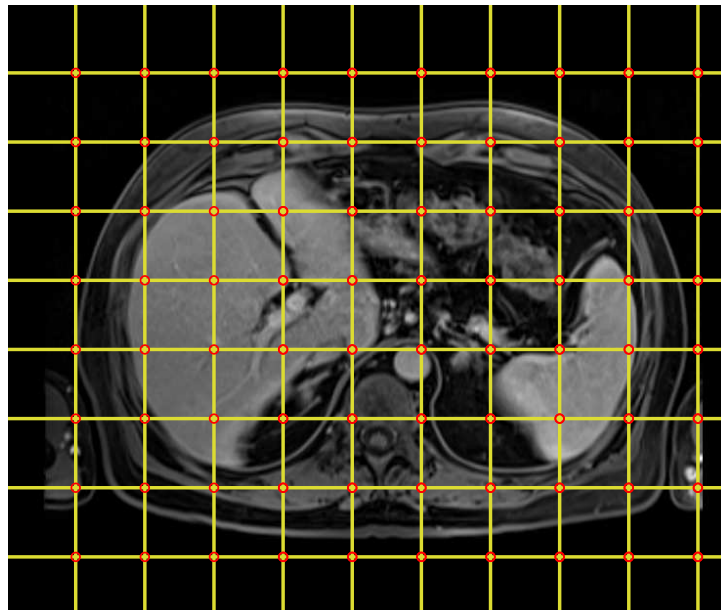


Abbildung 2.12.: Bei der B-Spline-Registrierung wird auf das Referenzbild ein Gitter gelegt, dessen Kontrollpunkte als Parameter fungieren.

gegeben, wobei  $\eta$  ein freier Steifigkeitsparameter ist und die Koeffizienten  $a, b, c, d, f_i$  durch ein lineares Gleichungssystem bestimmt werden, welches im Wesentlichen  $\mathbf{d}(\mathbf{r}_i) = \mathbf{d}_i$  garantiert.

Die robuste automatische Detektion von korrespondierenden Punktpaaren, die noch dazu alle Bildbereiche abdecken, ist insbesondere bei multimodalen Bildpaaren nahezu unmöglich. Daher werden oft die Kontrollpunkte  $\mathbf{r}_i$  gleichmäßig auf einem Gitter im Referenzbild verteilt (siehe Abbildung 2.12). Die Deformationsstützvektoren  $\mathbf{d}_i$  werden nun zu freien Parametern in einem  $3n$ -dimensionalen Optimierungsproblem einer bildbasierenden Fehlermetrik (vgl. Gleichung 2.3). Da, wie man in Gleichung 2.5 sieht, alle Kontrollpunkte einen Einfluss auf das gesamte Deformationsfeld haben, ist diese Optimierung für die in der Praxis benötigte Anzahl an Stützstellen nicht effizient. Eine Abhilfe schaffen kubische B-Splines [57], bei denen der Einfluss der Kontrollpunkte lokal begrenzt ist und daher auch für eine große Anzahl von Stützstellen geeignet sind.

Da aber auch hierbei aufgrund der hohen Anzahl an Parametern praktisch nur lokale Optimierungsverfahren in Frage kommen, ist die Capture Range sehr begrenzt. Um dieses bekannte Problem in den Griff zu bekommen kommen oft Multiresolution-Ansätze zum Einsatz. Ein weiteres, durch die große Flexibilität bedingtes Problem ist, dass es, wenn keine weiteren Einschränkungen getroffen werden, zu degenerierten Deformationsfeldern kommen kann. Strukturen in dem Referenzbild können auf die falschen Strukturen im Modellbild abgebildet werden, die Größe von einzelnen Strukturen kann sich drastisch verändern und es kann sogar zu Überschneidungen kommen.

**Komplett freie Registrierung:** Der höchstmögliche Grad an Freiheit besteht darin, jedem einzelnen der  $N$  Voxel im Referenzbild einen eigenen freien Deformationsvektor  $\mathbf{d}_i \in \mathbb{R}^3, i \in [1 \dots N]$  zuzuordnen und dann für jeden Voxel die optimale Deformation zu bestimmen. Dies geschieht z.B. in der 1998 von Thirion vorgestellten *Demons* Registrierung [58] durch ein Diffusionsmodell. In der

sog. Demons 1 Registrierung wird in Anlehnung an die Maxwellschen Dämonen für jeden Voxel  $i$  der dazugehörige Deformationsvektor  $\mathbf{d}_i$  iterativ verändert, so dass sich die Intensitäten in dem Referenzbild und dem transformierten Modellbild annähern. Dies funktioniert freilich nur auf Bildern derselben Aufnahmetechnik.

Mit diesem größtmöglichen Grad an Freiheit kann theoretisch jedes beliebige Deformationsfeld, also auch das perfekte, abgebildet werden. In der Praxis kann dies allerdings nur selten tatsächlich gefunden werden, da, wie oben bereits erwähnt, mit zunehmender Freiheit auch die Störanfälligkeit steigt. Um dennoch Transformationen mit hohen Freiheitsgraden verwenden zu können, ist *Regularisierung* nötig, das heißt, es muss Vorwissen über die möglichen Transformationen in den Registrierungsprozess eingebracht werden, so dass nicht plausible Transformationen (bspw. Überdeckungen, Unstetigkeiten, Volumenänderungen) ausgeschlossen bzw. bestraft werden.

#### 2.2.2.2. Vorwissen

Eine intuitiv einleuchtende Annahme, die über ein Deformationsfeld getroffen werden kann, ist, dass es glatt sein soll, d.h. dass  $\mathbf{x} \approx \mathbf{y} \Rightarrow \mathbf{d}(\mathbf{x}) \approx \mathbf{d}(\mathbf{y})$  gelten soll. Dies ist bei den Spline-basierten Transformationen automatisch gegeben. Bei der Demons-Registrierung wird dies dadurch erreicht, dass ein Glättungsfilter nach jeder Iteration auf das Deformationsfeld angewendet wird.

Die am häufigsten verwendete Art der Regularisierung besteht allerdings darin, die Zielfunktion der Optimierung (Gleichung 2.3) um einen Regularisierungs- oder Plausibilitätsterm  $P(\mathbf{p}) : \mathbb{R}^n \mapsto \mathbb{R}_0^+$ , der unwahrscheinliche Transformationen bestraft, folgendermaßen zu erweitern:

$$E(\mathbf{p}) = F(R, M_T(\mathbf{p})) + \lambda \cdot P(\mathbf{p}) , \quad (2.6)$$

wobei  $\lambda$  den Einfluss der beiden Terme steuert. Man beachte, dass  $P$  nur von der Transformation abhängt und damit einer internen Energie gleichkommt. Neben dem Plausibilitätsterm ist die Regularisierung vor allem aber von der verwendeten Transformationsklasse gegeben, wobei ersteres eine weiche Randbedingung und letzteres hingegen eine harte Randbedingung darstellt.

Die Art der Regularisierung kann durch ein Kontinuum beschrieben werden, an dessen einem Ende *Lokale Regularisierung* und an dem anderen Ende die *Globale Regularisierung* steht. Auf diese Art und Weise kann das in Abbildung 2.11 dargestellte Transformationsklassenkontinuum um die Dimension des Vorwissens erweitert werden (siehe Abbildung 2.13).

**Lokale Regularisierung:** Bei der lokalen Regularisierung fließt lediglich örtlich begrenztes Vorwissen über die zu erwartende Transformation ein. Dies kann beispielsweise die zuvor erwähnte Glattheitsanforderung sein, welche implizit durch bspw. die Verwendung von Splines und zusätzlich explizit durch einen entsprechenden Plausibilitätsterm modelliert werden. Ein von Wahba et al. [59] vorgestellter und danach unter vielen anderen auch in [57] verwendeter Glattheitsplausibilitätsterm sieht folgendermaßen aus:

$$P_{\text{glattheit}} = \frac{1}{V} \iiint_V \left[ \left( \frac{\partial^2 T}{\partial x^2} \right)^2 + \left( \frac{\partial^2 T}{\partial y^2} \right)^2 + \left( \frac{\partial^2 T}{\partial z^2} \right)^2 + 2 \left( \frac{\partial^2 T}{\partial xy} \right)^2 + 2 \left( \frac{\partial^2 T}{\partial xz} \right)^2 + 2 \left( \frac{\partial^2 T}{\partial yz} \right)^2 \right] dx dy dz . \quad (2.7)$$

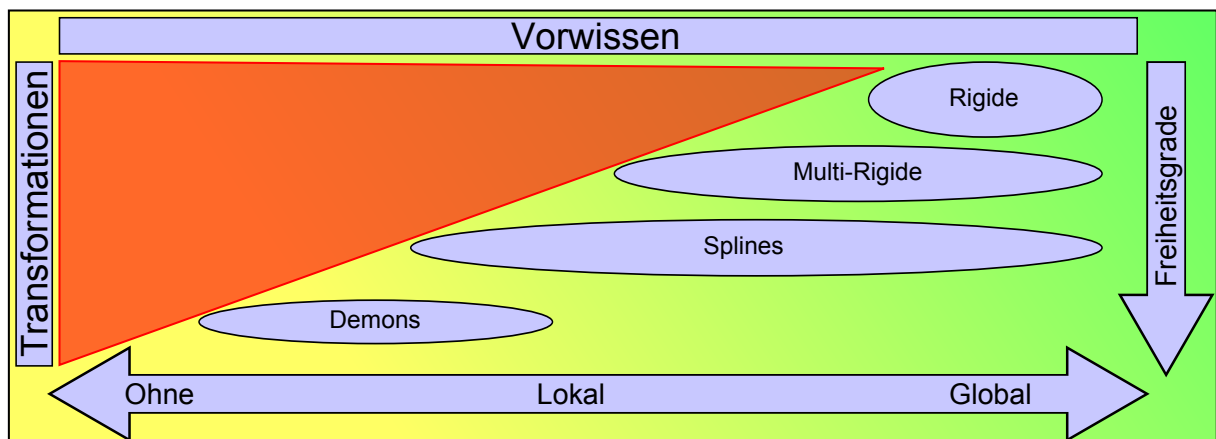


Abbildung 2.13.: Die Einteilung der Transformationen wird um die Dimension des Vorwissens erweitert. Die Robustheit der Registrierungsverfahren nimmt vom gelben hin zum grünen Bereich diagonal zu. Der rot markierte Bereich kann durch die harte Randbedingung der Art der Transformation nicht abgedeckt werden.

Das zugrundeliegende physikalische Modell ist dasselbe wie das durch Gleichung 2.5 beschriebene der Thin-Plate-Splines.

Abgesehen von der Glattheit, sind weitere lokale Plausibilitätssterme möglich. Rohlfing et al. schlagen in [60] einen lokalen Volumenerhaltungsterm vor, der darauf basiert, dass die Determinante der Jacobimatrix der Transformation den Wert eins hat, wenn weder eine Stauchung, noch eine Spreizung vorliegt. In [61] wird dies von Loeckx et al. erweitert, so dass nicht nur Volumenerhaltung, sondern lokale Starrheit gefordert wird.

In [62] und [63] werden ebenfalls lokale Starrheit gefordert, allerdings nur in einigen (zuvor identifizierten) Bereichen. Dadurch, dass diese Bereiche (z.B. Läsionen oder Knochen) zuvor identifiziert werden müssen, verliert diese Art der Regularisierung zum Teil ihren lokalen Charakter.

**Globale Regularisierung:** Bei der globalen Regularisierung hingegen fließt Wissen über die gesamte Transformation ein. Dies kann z.B. sein, dass sie an jeder Stelle dieselbe sein soll, oder dass an bestimmten Orten nur bestimmte Transformationen vorkommen können bzw. wahrscheinlicher sind. Ersteres ist bei allen linearen Abbildungen, wie z.B. die Ähnlichkeitstransformation (Rotation + Translation + isotrope Skalierung), automatisch gegeben. Zusätzlich kann dennoch ein globaler Plausibilitätsterm zur Anwendung kommen, der bspw. bestimmte erwartete Skalierungen bevorzugt. Je globaler eine Regularisierung ist, desto höher ist der Spezialisierungsgrad einer Methode, weil die Menge der möglichen Transformationen eingeschränkt wird. Gleichzeitig steigt dadurch auch die Robustheit der Methode gegenüber Störeinflüssen der Bilder.

### 2.2.3. Regmentierung

Obwohl die Zielsetzung von Segmentierung und Registrierung grundverschieden ist, ist die Methodik der modellbasierten Segmentierungsverfahren praktisch identisch mit Registrierungsverfahren. In bei-

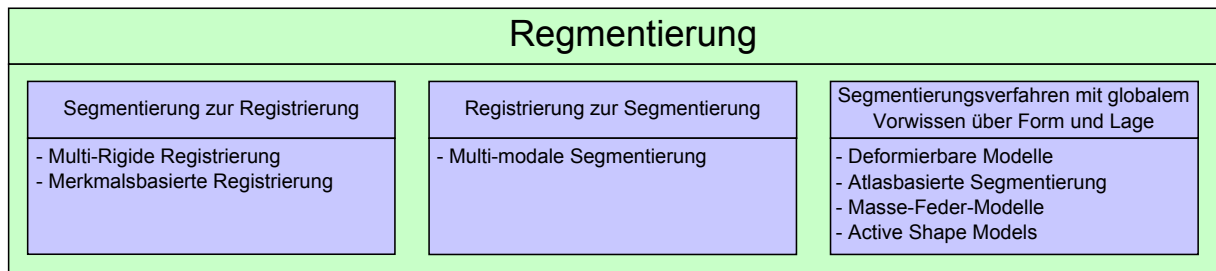


Abbildung 2.14.: Die drei verschiedenen Arten der Regmentierung mit Anwendungsbeispielen

den Fällen wird dasselbe Energiefunktional (vgl. Gleichung 2.1 und 2.6) minimiert, welches aus einem Datenterm und einem Regulierungsterm besteht. Das transformierte Modell kann sich, muss sich jedoch nicht, dabei unterscheiden. Während bei Verfahren, die zur Segmentierung eingesetzt werden, häufig Polygonnetze zum Einsatz kommen, dient zur Registrierung meist eins der beteiligten Bilder als Modell. Die atlasbasierten Verfahren treiben dies auf die Spitze, indem in einem zur Segmentierung verwendeten Verfahren ebenfalls ein Bild als Modell verwendet wird, wobei zu diesem Bild vorab semantische Information über die Objektzugehörigkeit vorhanden sein muss. Aufgrund dieser Gemeinsamkeiten haben mein Kollege Marius Erdt und ich in [5] diese Klasse von Bildverarbeitungsverfahren als *Regmentierung* bezeichnet.

Neben bildbasierten Registrierungsverfahren und modellbasierten sowie lageorientierten Segmentierungsverfahren, gehören dazu auch Verfahren, die zunächst segmentieren, um anschließend zu registrieren, oder registrieren, um anschließend zu segmentieren (siehe Abbildung 2.14). Ein Beispiel für ersteres sind Registrierungsverfahren, die unterschiedliche Regionen des Bildes auf unterschiedliche Art und Weise registrieren. Die Identifikation dieser Regionen erfolgt dabei mittels eines Segmentierungsverfahrens. Bspw. muss in einem multi-rigiden Verfahren zur Registrierung [54] die Segmentierung der einzelnen rigiden Regionen (Knochen) bekannt sein, bevor diese dann registriert werden können und daraus die Deformation für das Weichteilgewebe bestimmt werden kann. In die letzte Kategorie fallen alle Verfahren zur Segmentierung, die die komplementäre Information in unterschiedlichen Bildmodalitäten derselben Szene ausnutzen. Die zugrunde liegenden multimodalen Bilder müssen dazu natürlich zueinander registriert vorliegen. In einigen Fällen laufen diese beiden Prozesse sogar gleichzeitig ab.





## 3. Stand der Forschung

### 3.1. Segmentierung von Lymphknoten aus CT-Datensätzen

Unter Segmentierung von Lymphknoten im Kopf-Hals-Bereich können zwei fundamental unterschiedliche Vorgänge verstanden werden. Zum einen ist das die Bestimmung der Lymphknotenregionen, die in Kapitel 2.1.1 vorgestellt wurden. Dabei sind die Objektgrenzen einzelner Lymphknoten nicht von Bedeutung. Das Hauptanwendungsgebiet ist die Bestimmung der Zielorgane zur Strahlentherapieplanung. Zum anderen versteht man unter Lymphknotensegmentierung die Segmentierung einzelner, relevanter Lymphknoten, die entweder durch ein automatisches Lymphknoten-Detektionsverfahren wie z.B. [64][65][66][67] oder durch einen Anwender zuvor ausgewählt wurden. Dies wird, wie schon in Kapitel 1.1 erklärt, zur Stadienbestimmung, Verlaufskontrolle und Merkmalsextraktion benötigt. Da der Fokus dieser Arbeit auf der Segmentierung einzelner Lymphknoten aus CT-Datensätzen liegt, wird in diesem Kapitel hierauf umfangreich eingegangen. Verfahren zur Segmentierung von Lymphknoten aus anderen Modalitäten sowie Verfahren zur Detektion von Lymphknotenregionen werden nur kurz angerissen.

#### 3.1.1. Lageorientierte Verfahren zur Lymphknotenregionensegmentierung

Da Lymphknoten an vielen verschiedenen Regionen im ganzen Körper vorkommen und das umliegende Gewebe sehr unterschiedlich sein kann, ist es bei der Segmentierung einzelner Lymphknoten nicht sinnvoll, Wissen über die relative Lage zu anderen Objekten einzubeziehen. Daher werden Lage-orientierte Verfahren nur zur Segmentierung ganzer Lymphknotenregionen eingesetzt. Hierbei wiederum ist die Einbeziehung der relativen Lage zu anderen Objekten zwingend erforderlich, weil die Lymphknotenregionen keinen abgeschlossenen Objekten entsprechen, sondern lediglich implizit über umliegende Landmarken definiert sind.

Commowick et al. [68] verwenden zur Segmentierung der Lymphknotenregionen II,III und IV einen Durchschnittsatlas, welcher aus 45 manuell konturierten CT-Datensätzen des Kopf-Hals-Bereichs erstellt wurde. Der Erstellungsprozess beinhaltet drei Schritte: Durchschnittsbilderstellung durch Registrierung und Mittelung der CT-Aufnahmen, Durchschnittssegmentierungserstellung durch den STAPLE-Algorithmus [69] und Symmetrisierung an der mittleren Sagittalebene. Die Anpassung des Atlas erfolgt mittels eines deformierbaren Registrierungsverfahrens, welches in Abschnitt 3.2.1 kurz beschrieben ist.

Gorthi et al. [70] verwenden ebenfalls einen atlasbasierten Ansatz, um die Lymphknotenregionen im Kopf-Hals-Bereich zu segmentieren (siehe Abbildung 3.1). Bei diesem Atlas handelt es sich lediglich um einen Datensatz, für den neben den Lymphknotenregionen auch andere Strukturen – nämlich die Außenkontur, Unterkiefer, Wirbel und die Trachea – manuell konturiert wurden. Diese zusätzlichen Strukturen werden in einem komplett frei deformierbarem Registrierungsprozess (vgl.

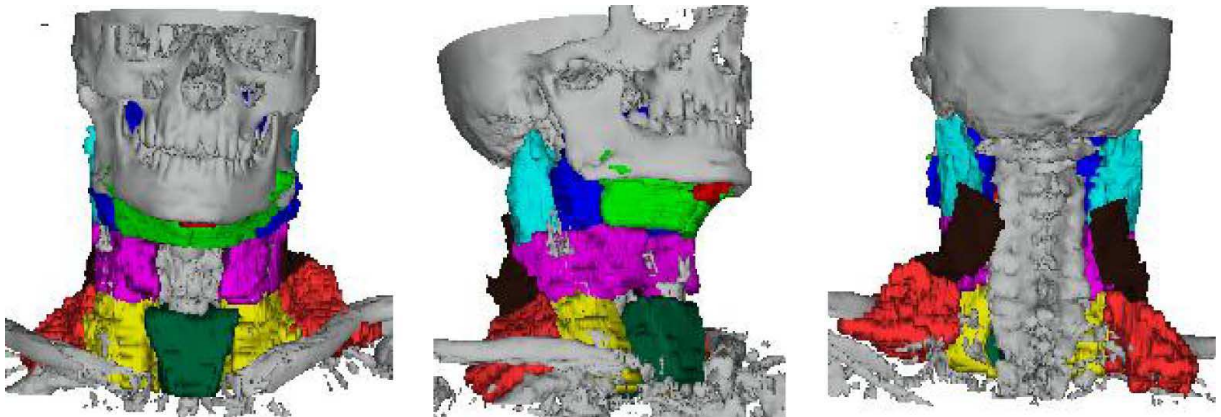


Abbildung 3.1.: Automatische Segmentierung der Lymphknotenregionen durch ein atlasbasiertes Segmentierungsverfahren (Abbildung aus [70])

Abschnitt 3.2.1) berücksichtigt. Wie in atlasbasierten Segmentierungsverfahren üblich werden die Lymphknotenregionen abschließend auf das Referenzbild übertragen.

Auch bei Teng et al. [71] kommt ein atlasbasiertes Verfahren zum Einsatz, wobei hier mehrere manuell konturierte CT-Datensätze als Atlas zur Verfügung stehen, von denen der geeignetste automatisch ausgewählt wird. Zunächst werden Landmarken, wie z.B. das Zungenbein, der Unterkiefer und die Halswirbelsäule, durch ein heuristisches Verfahren, welches ein 2D-Schwellwertverfahren mit einem 3D-deformierbaren Modell kombiniert, grob segmentiert. Welcher Atlas als Modellbild verwendet wird, ergibt sich durch einen Vergleich eines Merkmalsvektors, welcher sowohl aus dem Modellbild als auch aus dem Referenzbild extrahiert wird. Dieser enthält das Gesamtvolumen der Kopf-Hals-Region, Oberflächen einiger Strukturen und Vektoren zwischen den Mittelpunkten von Landmarken. Der passendste Atlas wird dann mit einer *Mutual Information* [72][73] basierten B-Spline-Registrierung an das Referenzbild angepasst (vgl. Abschnitt 3.2.1). Abschließend werden die Lymphknotenregionen auf das Referenzbild übertragen.

#### 3.1.2. Vergleichende Studien

Wie in den folgenden Abschnitten berichtet wird, wurden einige Methoden speziell auf das schwierige Problem der Segmentierung von einzelnen Lymphknoten zugeschnitten. Ein Vergleich dieser Methoden untereinander fehlt allerdings weitestgehend. Lediglich in [74] und [75] werden allgemeine Segmentierungsmethoden hinsichtlich ihrer Eignung zur Segmentierung einzelner Lymphknoten aus CT- bzw. MRT-Aufnahmen einander gegenübergestellt.

Die wahrscheinlich früheste Arbeit zur Segmentierung einzelner Lymphknoten aus CT-Datensätzen wurde im Jahre 1996 von Rogowska et al. [74] vorgestellt. Darin wurde ein *automatisches Schwellwertverfahren* (AS), ein *semi-automatisches lokales Schwellwertverfahren* (LS), die *Wasserscheiden-transformation* (WT) und *Snakes* (SN) auf deren Eignung zur Anwendung auf Lymphknoten in einzelnen zweidimensionalen CT-Schichten überprüft. Bei allen Verfahren, mit Ausnahme von AS, ist eine manuelle Initialisierung erforderlich, die deutlich über eine einfache Auswahl des Lymphknotens hinaus geht (z.B. grobes Einzeichnen der Kontur). Die Evaluierung fand auf Phantomdaten



*Größeninvarianz* unabhängig von der Bildmodalität eine wichtige Anforderung an ein Lymphknoten-segmentierungsverfahren.

#### 3.1.3. Grauwert-, Regionen- und Kanten-orientierte Verfahren

Wie bereits geschildert haben Grauwert-, Regionen- und Kanten-orientierte Verfahren keine große Aussicht auf Erfolg bei der Segmentierung von einzelnen Lymphknoten aus CT-Aufnahmen. Aufgrund der Einfachheit wird so ein Verfahren, nämlich das *Region-Growing*, dennoch als Teilschritt des Lymphknoten-Detektions-Algorithmus von Kitasaka et al. [64] verwendet. In einem späteren Schritt werden die entstandenen Regionen in Bezug auf Form und Größe analysiert. Nicht passende Regionen werden aussortiert. Evaluert wurde das Verfahren allerdings nur in puncto Detektionssensitivität und -spezifität, nicht jedoch in puncto Segmentierungsgenauigkeit.

#### 3.1.4. Verfahren mit globalem Formwissen

Obwohl Lymphknoten, wie in Abschnitt 2.1.1 erläutert, eine große Formvarianz aufweisen, werden in der Literatur dennoch Verfahren, welche globales Formwissen einbeziehen, vorgestellt. Dieses Wissen wird aber nie in Form eines statistischen Formmodells eingebracht, wie es bei vielen Organen, wie z.B. der Leber, üblich ist, weil für Lymphknoten kein charakteristischer Formenraum (z.B. Ausbuchtung immer oben links) existiert. Stattdessen wird auf einfache geometrische Formen, wie *Ellipsoid* und *Kugel* zurückgegriffen, von denen allerdings relativ große Abweichungen erlaubt sind.

**Ellipsoid:** Unal et al. [76] verwenden eine Ellipse als geometrisches Modell zur gleichzeitigen Segmentierung und Registrierung von Lymphknoten aus 2D-MRT-Schichten von  $n$  unterschiedlichen Sequenzen. Parametrisierung der Ellipse und der rigid Transformationen führt zu einem  $(5 + 4n)$ -dimensionalen Optimierungsproblem, welches ein Regionen- und einen Kanten-basiertes Energiefunktional minimiert. Die Initialisierung erfolgt durch ein Rechteck. Um ein 3D-Volumen zu erhalten, wird das 2D-Ergebnis in benachbarte Schichten propagiert. Die Evaluierung erfolgte durch den Vergleich mit manueller Konturierung auf 30 Lymphknoten. Durchschnittlich wurde 11.00/13.05/19.96 % des Lymphknotens auf  $T2/T2^{*1}/T2^{*2}$ -gewichteten MRT-Aufnahmen nicht in die Segmentierung einbezogen. Zhang et al. [77] bringen die Distanz zu einer Ellipse als Formwissen in einem Graph-Cut basierten Segmentierungsverfahren von Lymphknoten aus 2D-Ultraschall-Schichten ein. Die Initialisierung erfolgt ebenfalls durch ein Rechteck, aus welchem auch die für den Graph-Cut benötigten initialen Objekt- und Hintergrundmarker extrahiert werden. Zusätzlich zu den üblichen auf Intensitätsunterschieden basierenden Kantengewichten  $w(p, q)$  benachbarter Pixel  $p, q$  und auf Intensitätswahrscheinlichkeiten basierenden Kantengewichten zu Quelle  $w(p, Q)$  und Senke  $w(p, S)$  wird die Berechnung der Kantengewichte dahingehend erweitert, dass eine große Distanz  $D$  zu einer Ellipse (siehe Abbildung 3.3(c)) zu einer geringeren Kantenwahrscheinlichkeit führt. Das mit Graph-Cut zu minimierende Gesamtenergiefunktional ergibt sich damit zu:

$$E = (1 - \mu) \left[ \lambda \sum_{\forall p} w(p, Q) + w(p, S) + \sum_{\forall \{p, q\} \in N} w(p, q) \right] + \mu \sum_{\forall \{p, q\} \in N} D \left( \frac{p + q}{2} \right), \quad (3.1)$$

wobei  $\lambda$  und  $\mu$  den Einfluss der einzelnen Terme regeln. In einem ersten Schritt wird der traditionelle Graph-Cut ohne globales Formwissen angewendet (siehe Abbildung 3.3(a)), aus welchem initiale

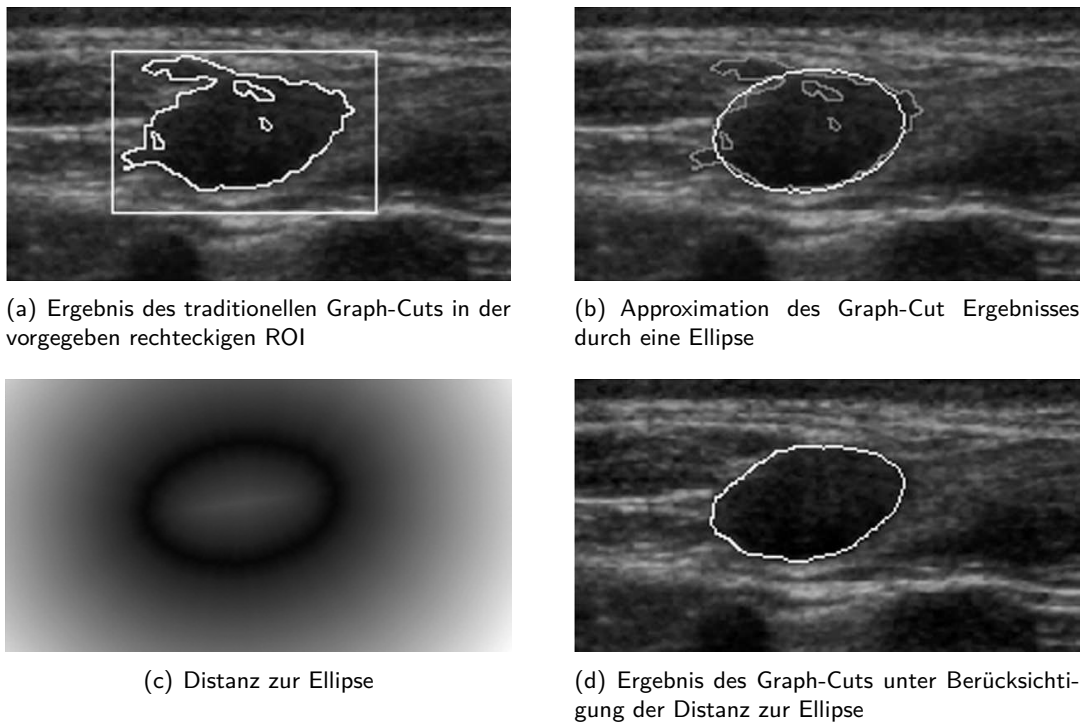


Abbildung 3.3.: Graph-Cut mit Ellipse als globales Formwissen zur Segmentierung von Lymphknoten aus Ultraschallbildern (Abbildungen aus [77])

Parameter einer Ellipse geschätzt werden (siehe Abbildung 3.3(b)). In weiteren Iterationen wechseln sich Graph-Cut und Ellipsenberechnung bis zur Konvergenz ab. Evaluiert wurde die Methode auf klinischen Daten von 60 Patienten. Mit einer durchschnittlichen *Hausdorff-Distanz* (HD) von 10.1 Pixeln hat das Einbringen von Formwissen mittels einer Ellipse einen großen Vorteil gegenüber dem traditionellen Graph-Cut (HD: 29.2) und liefert zu manueller Segmentierung vergleichbare Ergebnisse (HD: 10.6).

Moltz et al. [78] stellten ein 3D-Verfahren vor, welches neben Lymphknoten auch leicht modifiziert auf Lebermetastasen und Lungenherde in CT-Aufnahmen angewendet wurde. Die Initialisierung erfolgt durch das manuelle Einzeichnen eines Strichs, der quer durch das zu segmentierende Objekt geht und so eine ROI definiert. Die Segmentierung erfolgt dann in mehreren Schritten (siehe Abbildung 3.4): i) Region-Growing um den Mittelpunkt der ROI; ii) Radiale Strahlen werden in alle Richtungen geschickt, bis diese entweder an die Region-Growing-Grenze stoßen oder aber die ROI verlassen; iii) Aus den Endpunkten aller Strahlen, die die ROI nicht verlassen, wird eine Ellipsoid so bestimmt, dass alle Endpunkte gerade so enthalten sind; iv) Wasserscheidentransformation auf dem Distanzbild des Region-Growing-Ergebnisses, wobei lokale Distanzbildmaxima innerhalb des Ellipsoids als Objektmarker und Maxima außerhalb des Ellipsoids sowie Punkte am Rand der ROI als Hintergrundmarker verwendet werden; v) Nachverarbeitung durch morphologische Operatoren. Die Evaluierung erfolgte auf 222 Lymphknoten im Brust-, Abdomen- und Beckenbereich. Allerdings wurden die Segmentierungsergebnisse nicht mit Referenzsegmentierungen verglichen, sondern lediglich durch zwei unabhängige Experten bewertet. In 78% der Fälle wurde die Segmentierungsqualität mit *gut* oder *sehr gut* bewertet.

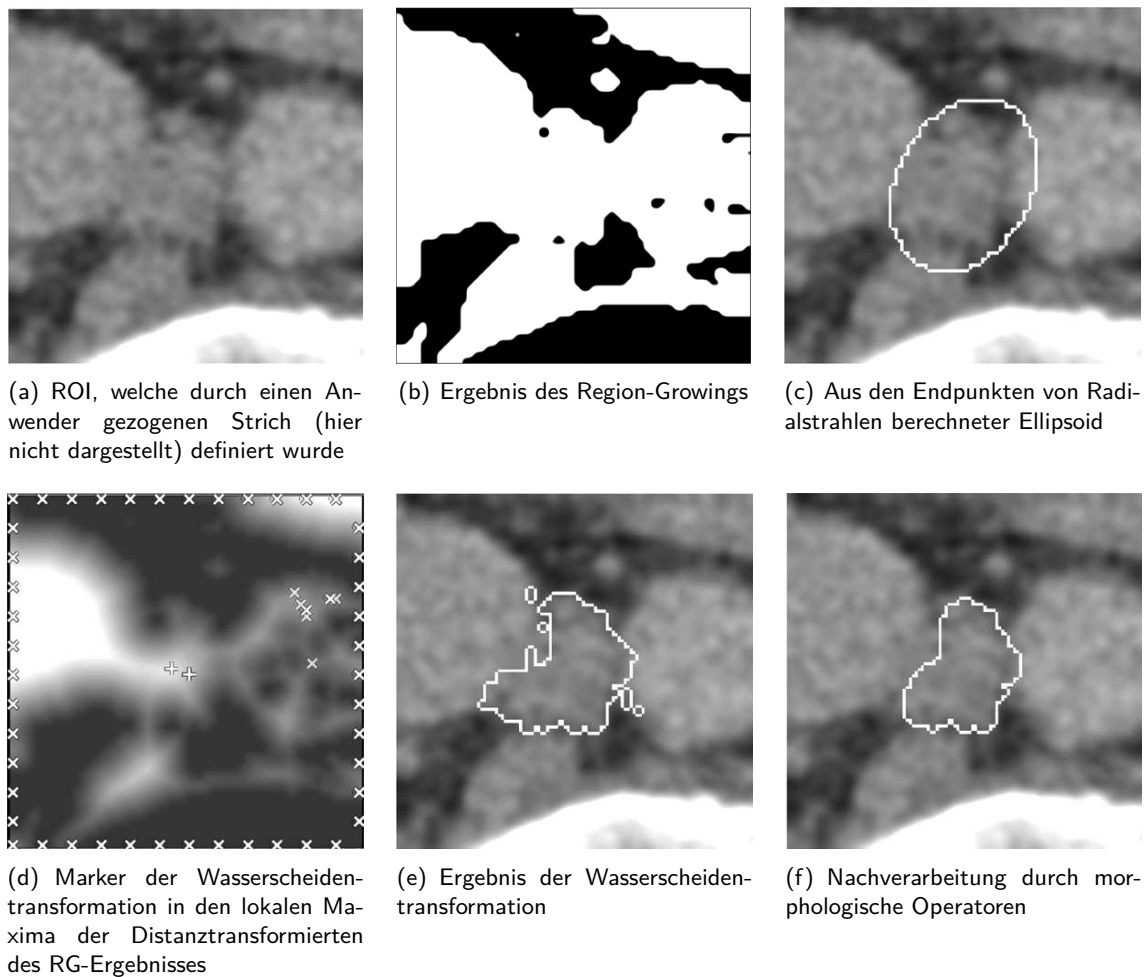
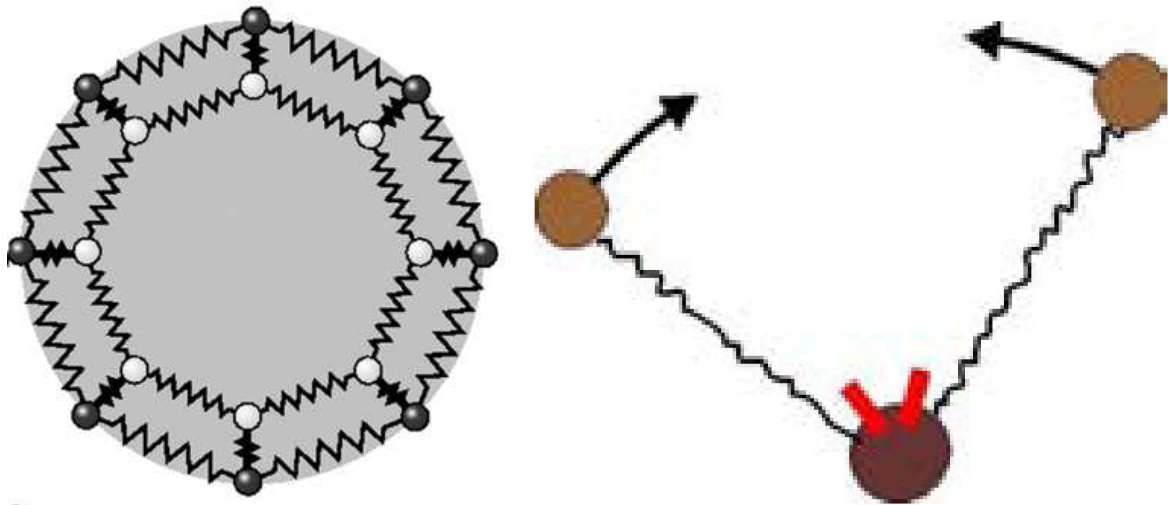


Abbildung 3.4.: Auf der Wasserscheidentransformation basierendes heuristisches Verfahren zur Segmentierung von Lymphknoten aus CT-Datensätzen (Abbildungen aus [78])

Maleike et al. [79] verwenden zur 3D-Segmentierung von Lymphknoten aus CT-Datensätzen ein deformierbares Modell, welches einem Ellipsoiden entspricht. Da das umliegende Gewebe von Lymphknoten sehr unterschiedlich ist (vgl. Abschnitt 2.1.1), kann dessen Aussehen nicht aus Referenzbildern gelernt werden. Stattdessen wird die Kostenfunktion so gewählt, dass für Lymphknoten typische Grauwerte innerhalb des Modells und andere Grauwerte außerhalb des Modells zu geringen Kosten führen. Darüber hinaus werden große Bildgradientenstärken an der Modelloberfläche bevorzugt. Initialisiert wird das 3D-Modell durch eine manuell eingezeichnete grobe 2D-Kontur in einer Schicht. Zusätzlich zu der automatischen Segmentierung wird dem Anwender anschließend auch eine manuelle Korrektur erlaubt. Dieses Verfahren wurde auf 29 Lymphknoten angewendet und mit einer Referenzsegmentierung verglichen. Es stellte sich heraus, dass das Segmentierungsverfahren auch ohne Korrektur bei einigen Fehlermaßen vergleichbare Ergebnisse wie die manuelle Segmentierung erzielt (Median der Hausdorff-Distanz: 2.7 vs. 2.6 mm, des Volumenüberlappungsfehlers: 0.19 vs. 0.21). Bei anderen Fehlermaßen schneidet das Verfahren allerdings deutlich schlechter ab (Median der durchschnittlichen Oberflächendistanz: 0.6 vs. 0.3 mm, der Volumendifferenz: 10.1 vs. 4.2%).



(a) Anordnung der Massepunkte mit Intensitätssensoren auf dem inneren Ring und Bildgradientensensoren auf dem äußeren Ring (b) Zur Erhaltung der initialen Winkel (rote Balken) werden Torsionskräfte auf die Massepunkte ausgeübt

Abbildung 3.5.: Masse-Feder-Modell zur Segmentierung von Lymphknoten aus CT-Datensätzen (Abbildungen aus [80] bzw. [47])

**Kugel mit fester Größe:** Anstatt eines Ellipsoiden verwenden Dornheim et al. [81][80] eine noch einfachere Form, nämlich eine Kugel, als Formwissen für ein stabiles Masse-Feder-Modell [47] (*Stable-Mass-Spring-Model*(SMSM)). Ein Massen-Feder-Modell besteht aus Massenpunkten  $i$ , auf die durch das Bild getriebene Kräfte  $\mathbf{F}_i$  wirken. Form- und Größenerhalt wird über Federn erreicht, die die Massepunkte  $i$  und  $j$  miteinander verbinden und *Federkräfte*  $\mathbf{F}_{i,j}$  ausüben. Das Modell passt sich nach den physikalischen Gesetzen der Newton-Mechanik in einem zeitdiskreten Prozess an ein Objekt an:

$$\mathbf{v}_{i,t+\Delta t} = \left( \mathbf{v}_i + \frac{w_f \cdot \sum_j \mathbf{F}_{i,j} + w_s(i) \mathbf{F}_i}{m_i} \cdot \Delta t \right) \cdot (1 - d), \quad (3.2)$$

wobei  $\Delta t$  dem Zeitschritt,  $w_s$  und  $w_f$  Gewichten der Kräfte,  $m_i$  der Masse des Massenpunktes  $i$  und  $d$  einem Dämpfungsfaktor entspricht. Zusätzlich zu den Federkräften kommen bei dem SMSM auch noch *Torsionskräfte* zum Einsatz, welche die Winkel zwischen den Federn erhalten (siehe Abbildung 3.5(b)). Die Kombination aus Feder- und Torsionskräften führt dazu, dass das Modell immer in die Ursprungsform zurückfällt, falls keine externen Kräfte wirken. Dadurch kann man von globalem Formwissen sprechen. Zur Segmentierung von Lymphknoten aus CT-Datensätzen werden die Massepunkte wie in Abbildung 3.5(a) angeordnet und um einen durch den Anwender vorgegebenen Saatpunkt initial platziert. Die Größe wurde dabei so klein gewählt (0.3 mm), dass sich das gesamte Modell innerhalb des Objekts befindet. Um eine Änderung der Größe zuzulassen, werden die Federgewichte des äußeren und inneren Rings sehr klein im Vergleich zu denen zwischen den Ringen gewählt. Die Massenpunkte des äußeren Rings werden durch *Gradientensensoren* in Regionen mit hoher Gradientenstärke in Richtung der Modellnormalen gedrückt. Zusätzlich werden *Intensitätssensoren* auf dem inneren Modellring platziert, welche diese Massepunkte in Regionen mit für Lymphknoten typischen CT-Intensitäten drücken. Um auch die Segmentierung von Lymphknoten mit zentralem nekrotischen Gewebe zu unterstützen, sind diese Intensitätssensoren allerdings nicht im ge-

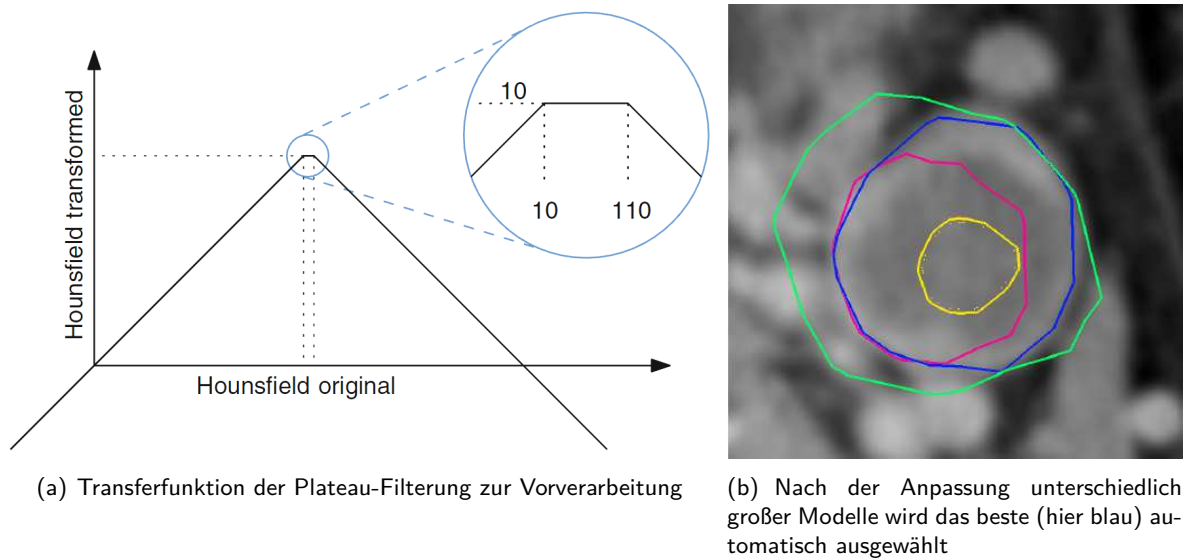
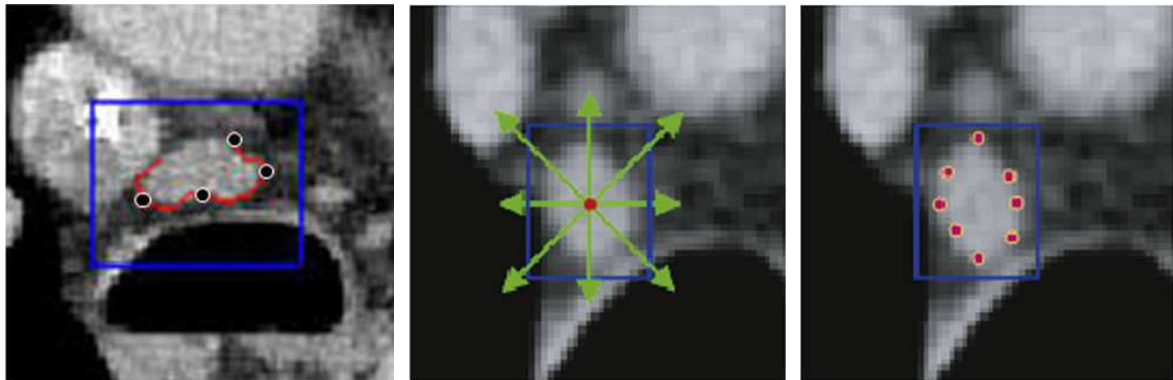


Abbildung 3.6.: Erweiterungen des Masse-Feder-Modells zur Erhöhung der Größeninvarianz (Abbildungen aus [82])

samen Modell verteilt. Eine quantitative Evaluation erfolgte auf 40 Lymphknoten, wobei hierbei das Modell durch zwei Punkte initialisiert wurde, welche die Modellgröße vorgaben. Die durchschnittliche Oberflächendistanz variierte zwischen 0.3 und 0.6 mm (Hausdorffdistanz: 1.7 – 3.9 mm) und war damit vergleichbar mit manueller Segmentierung.

Um auch pathologische Lymphknoten erfolgreich aus CT-Datensätzen segmentieren zu können, stellten Dornheim et al. in [82] und [83] zwei Erweiterungen zu dem SMSM-Segmentierungsverfahren vor. Das Bild wird zunächst einer sog. *Plateau-Filterung* unterzogen, bevor die Sensoren darauf zugreifen (siehe Abbildung 3.6(a)). Dadurch werden alle Bildintensitäten so transformiert, dass i) evtl. vorhandene Bildgradienten innerhalb der für Lymphknoten üblichen Grauwerte die Intensitätssensoren nicht stören, ii) die Gradientenstärken erhalten bleiben und iii) Lymphknoten nun die Strukturen mit den hellsten Intensitäten sind. Letzteres hat zur Folge, dass die Gradientensensoren durch die Einbeziehung der Gradientenrichtung stärker auf Lymphknotengrenzen reagieren, nicht jedoch auf Bildgradienten benachbarter Strukturen mit eigentlich höheren Bildintensitäten, wie z.B. kontrastverstärkte Gefäße. Ferner wird eine *Multi-Modell-Segmentierung* angewendet (siehe Abbildung 3.6(b)). Das bedeutet, dass Modelle mit unterschiedlichen Parametern (hier initiale Modellgröße) auf den selben Datensatz angepasst werden. Im Anschluss daran wird das am besten angepasste Modell durch eine *Anpassungsgütefunktion* [84] automatisch bestimmt, welche auf der Form des angepassten Modells und den Bildsensorwerten basiert. Das ausgewählte Modell bestimmt die endgültige Segmentierung, könnte darüber hinaus aber auch noch zur Charakterisierung des Lymphknotens eingesetzt werden. Die Evaluierung erfolgte durch Vergleich mit [81] auf 21 Lymphknoten. In Bezug auf durchschnittliche Oberflächendistanz und Hausdorff-Distanz wurde eine leichte Verbesserung erzielt (0.41 vs. 0.53 mm bzw. 3.15 vs. 3.51 mm).





(a) Interaktive Segmentierung durch manuelle Vorgabe einer ROI und Konturpunkten. (b) Automatische Detektion von Konturpunkten auf Radialstrahlen. (c) Automatisch detektierte Konturpunkte zur Verbindung mittels Live-Wire

Abbildung 3.7.: Schichtweise Segmentierung von Lymphknoten mit Live-Wires (Abbildungen aus [85])

### 3.1.5. Verfahren mit lokalem Formwissen

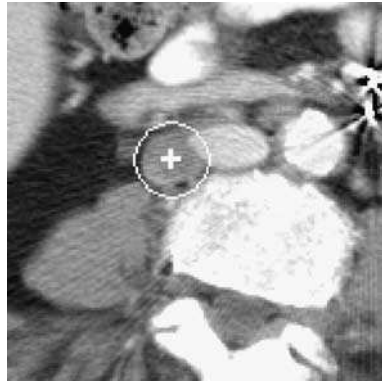
Aufgrund des geringen Kontrastes zu benachbarten Strukturen in CT-Aufnahmen ist die Verwendung von Formwissen vorteilhaft. Globales Formwissen setzt voraus, dass die gesamte zu erwartende Form in etwa bekannt ist. Da dies bei Lymphknoten allerdings nur sehr begrenzt der Fall ist, besteht ein oft verwendeter Ansatz darin, nur lokales Formwissen zu verwenden, also eine lokal glatte Kontur zu erzwingen, ohne dabei aber eine konkrete globale Form zu bevorzugen.

**Live-Wire:** Zur Segmentierung von Lymphknoten aus CT-Aufnahmen im Brustbereich stellen Lu et al. in [85] ein interaktives Verfahren vor, welches auf der *Live-Wire*- oder *Intelligent-Scissors* [86]-Segmentierung basiert. In einer repräsentativen 2D-Schicht bestimmt der Anwender zunächst eine ROI, die den Lymphknoten komplett enthält. Anschließend gibt er einige wenige Punkte auf der Kontur vor, die automatisch miteinander verbunden werden (siehe Abbildung 3.7(a)). Dabei wird aus der ROI ein Graph erstellt, bei dem jeder Pixel einem Knoten entspricht und für die Übergänge benachbarter Pixel lokale Übergangskosten, welche unter anderem die Bildgradientenstärke und -richtung einbeziehen, berechnet werden. Der Pfad mit den geringsten Kosten zwischen zwei Punkten wird durch den Algorithmus von Dijkstra bestimmt. Unter der Annahme, dass die Struktur in benachbarten Schichten sehr ähnlich aussieht, werden die Saatpunkte für die angrenzenden Schichten automatisch aus der aktuellen Kontur bestimmt. Dieser Vorgang wird solange fortgesetzt, bis eines von mehreren Abbruchkriterien erfüllt ist. Alternativ zur manuellen Vorgabe von Konturpunkten in der Referenzschicht können diese auch automatisch bestimmt werden. Ausgehend von einem einzigen vorgegebenen Saatpunkt innerhalb des Lymphknotens werden Radialstrahlen in alle Richtungen geschickt (siehe Abbildung 3.7(b)). Auf jedem dieser Strahlen fungiert der Punkt mit der höchsten Gradientenstärke als Konturpunkt für den folgenden Live-Wire-Segmentierungsprozess (siehe Abbildung 3.7(c)). Die Evaluierung erfolgte quantitativ auf 50 Lymphknoten. Im Vergleich zu einer Referenzsegmentierung erzielte das interaktive Verfahren einen durchschnittlichen DSC von  $0.81 \pm 0.07$ . Eine automatische Bestimmung der Konturpunkte führte zu dem leicht schlechteren

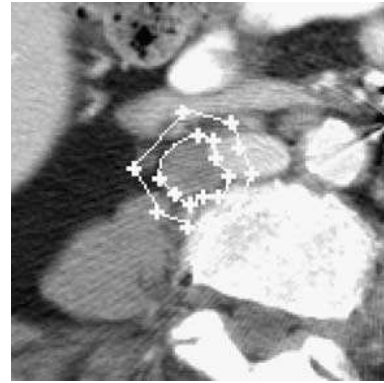
Wert  $0.79 \pm 0.08$ . Prinzipbedingt versagt die automatische Detektion der Kontrollpunkte bei Lymphknoten mit geringem Bildkontrast oder falls Strukturen mit hohem Bildkontrast (bspw. Knochen) sich in der Nähe des Lymphknotens befinden.

**Snakes und aktive Konturen:** 1997 verwendeten Honea et al. in einem der ersten vorgestellten Ansätze zur Lymphknotensegmentierung [87] einen *Snake* [39]-basierten Ansatz, welcher zunächst auf einer einzelnen repräsentativen 2D-Schicht zur Anwendung kommt. Die Initialisierung der Kontur erfolgt durch eine Ellipse, welche der Anwender durch das Einzeichnen von zwei orthogonalen Strichen definiert. Da so die Größe in etwa vorgegeben ist, führt eine lokale Anpassungsstrategie weder zu einem Kollaps der Kontur (Gefahr bei einer zu kleinen Initialisierung im homogenen Inneren, in dem keine geeigneten Punkte gefunden werden können) noch zu einer Anpassung an eine falsche Struktur (Gefahr bei einer zu großen Initialisierung). Die interne Energie ist durch den bei Snakes üblichen Stetigkeits- und Glattheitsterm gegeben. Die externe Energie für einen Punkt  $\mathbf{x}$  ist durch  $E_{\text{ext}}(\mathbf{x}) = h_{\text{rot}} \star I(\mathbf{x})$  gegeben, wobei  $I(\mathbf{x})$  dem Eingabebild und  $h_{\text{rot}}$  einem rotierten Prewitt-Operator zur Kantendetektion entspricht. Die Rotation ist durch die Verbindung des Schwerpunktes aller Konturpunkte mit dem Punkt  $\mathbf{x}$  gegeben, wodurch nur Kanten betrachtet werden, die von Innen nach Außen einen Hell-Dunkel-Verlauf aufweisen. Um die Anfälligkeit für starke Bildgradienten außerhalb des Lymphknotens zu reduzieren, erfolgt die Kandidatenpunktsuche in der ersten Iteration nur in das Innere der Kontur. In allen folgenden Iterationen wird dann in einer  $3 \times 3$ -Nachbarschaft gesucht. Aufgrund der hohen Varianz an Größe, Form und Bildcharakteristik einzelner Lymphknoten haben die Werte der einzelnen Energieterme für unterschiedliche Lymphknoten auch unterschiedliche Größenordnungen und eine sinnvolle a priori Gewichtung der Energieterme ist nicht sinnvoll. Daher findet vor der gewichteten Addition zur Gesamtenergie eine Normalisierung der einzelnen Energieterme statt. Um eine 3D-Segmentierung zu erhalten, wird das Segmentierungsergebnis einer Schicht als Initialisierung für die Segmentierung in den benachbarten Schichten verwendet. Die Methode wurde quantitativ auf 25 künstlichen Phantom-Lymphknoten durch Volumenmessung evaluiert. Die durchschnittliche Abweichung betrug dabei 4.5%. Zusätzlich wurde die Methode auch auf reale Lymphknoten angewendet. Dabei zeigte sich allerdings, dass die sukzessive 2D-Segmentierung oftmals nicht zu einer zufriedenstellenden 3D-Segmentierung führt, weil die Konturunterschiede in benachbarten Schichten so groß sein können, dass sich die Kontur an eine falsche Struktur anpasst.

Um dieses Problem zu umgehen, entwickelte dieselbe Gruppe zwei Jahre später in [88] ein echtes 3D-Segmentierungsverfahren, welches ein initial kugelförmiges Modell durch Energieminimierung an einen CT-Datensatz anpasst. Dieses Modell wird durch einen einzelnen Punkt platziert und ist so klein, dass es sich komplett innerhalb des Lymphknotens befindet. Expandiert das Modell, wird ein Punkt, der vom Schwerpunkt des Modells weit entfernt liegt durch drei Punkte ersetzt, um die Punktdichte auf der Oberfläche fein genug zu halten. Die auf *Simulated Annealing* basierende Anpassungsstrategie wählt in jeder Iteration einen einzelnen Modellpunkt zufällig aus, verschiebt diesen zufällig und behält die Verschiebung bei, wenn sich die Gesamtenergie verringert. Wird die Gesamtenergie erhöht, wird mit einer von dieser Energieerhöhung abhängigen Wahrscheinlichkeit der Punkt trotzdem beibehalten, so dass das Modell nicht zwangsläufig in lokalen Minima des Energiefunktionalen stecken bleibt. Um eine Expansion des Modells zu erreichen, wird bei der zufälligen Verschiebung eines Modellpunktes darauf geachtet, dass Verschiebungen aus dem Modell hinaus wahrscheinlicher sind. Daher wird bei dem Modell auch von einem *Balloon* gesprochen. Die interne Energie besteht aus zwei Termen, einer, der für eine gleichmäßige Verteilung der Modellpunkte auf der Oberfläche sorgt, und einer, der die



(a) Manuell gesetzte Abbruchbedingung für den Expansionsprozess in einer Schicht



(b) In allen weiteren Schichten dient das Segmentierungsergebnis der Nachbarschicht als Abbruchbedingung

Abbildung 3.8.: Level-Sets zur schichtweisen Segmentierung von Lymphknoten (Abbildungen aus [89])

Krümmung des Modells gering hält. Die externe Energie des Punktes  $\mathbf{x}$  ist durch

$$E_{\text{ext}} = \begin{cases} -|\nabla I(\mathbf{x})|^2 & , \text{ falls } \angle(\mathbf{r} - \nabla I(\mathbf{x})) < 30^\circ \\ -\cos(\mathbf{r} - \nabla I(\mathbf{x}))|\nabla I(\mathbf{x})|^2 & \text{sonst} \end{cases} \quad (3.3)$$

gegeben, wobei  $\nabla I(\mathbf{x})$  Bildgradientenvektor des Punktes  $\mathbf{x}$  ist und  $\mathbf{r}$  der Richtung vom Modellschwerpunkt zu  $\mathbf{x}$  entspricht. Diese Richtungsgewichtung erlaubt es, Hell-Dunkel-Verläufe zu bevorzugen und Bildgradienten, die orthogonal zu der Modellnormalen sind, zu ignorieren. Unglücklicherweise wurde die Methode nur auf künstlichen Daten evaluiert.

Yan et al. stellen in [89] ein *Level-Sets*-basiertes Segmentierungsverfahren für Lymphknoten aus CT-Aufnahmen vor. Eine sequenzielle Anwendung auf 2D-Schichten führt zu einer 3D-Segmentierung. Neben dem Bildgradienten werden bei der Anpassung auch Bildintensitäten berücksichtigt, wodurch Boundary-Leaking verhindert werden soll. Der Anwender bestimmt in einer repräsentativen 2D-Schicht neben einem Saatpunkt auch einen Kreis um diesen Punkt, dessen Radius so gewählt wird, dass der Lymphknoten gerade so komplett enthalten ist (siehe Abbildung 3.8(a)). Dieser Kreis dient als hartes Abbruchkriterium für den Anpassungsprozess. Für alle weitere Schichten dient die aus der Nachbarschicht bereits extrahierte Kontur als Abbruchkriterium (siehe Abbildung 3.8(b)). Ein Nachverarbeitungsschritt ermöglicht es dem Anwender das Segmentierungsergebnis anzupassen. Zur Evaluierung wurde das Verfahren auf ca. 400 Lymphknoten angewendet. Allerdings wurden keine quantitativen Ergebnisse vorgestellt.

Feulner et al. verwenden für den Segmentierungsteil ihres Lymphknotendetektions- und segmentierungsverfahren [90] einen Graph-Cut-basierten Ansatz. In einem Bereich um einen Saatpunkt wird aus dem Voxelgitter ein Graph erstellt. Die Kantenkapazitäten ergeben sich aus der Wahrscheinlichkeit der Objekt- und Objektgrenzenzugehörigkeit, welche aus Histogrammen bzw. gemeinsamen Histogrammen von annotierten Trainingsdaten gelernt wurden. Um dem sog. *small-cut*-Problem – also dem Fall, dass der Kapazitätsengpass sich direkt um die Quelle oder die Senke herum befindet – zu entgehen werden die Kantenkapazitäten mit dem Faktor  $\frac{1}{r}$  ( $r$  entspricht der Entfernung

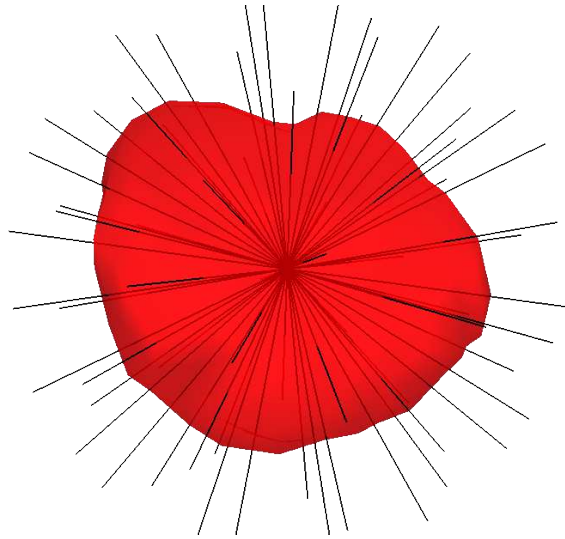


Abbildung 3.9.: Repräsentation einer Oberfläche durch Radialstrahlen.

eines Punktes zu dem Mittelpunkt – also dem Radius) gewichtet. Dies hat zur Folge, dass keine bestimmte Objektgröße bevorzugt wird und gleichzeitig rundere Strukturen aufgrund der geringeren Oberfläche bevorzugt werden. Das Verfahren wurde umfangreich in Bezug auf Detektionssensitivität und -spezifität evaluiert, nicht jedoch in Bezug auf Segmentierungsgenauigkeit.

**Radialstrahlen:** Eine weitere Möglichkeit, lokales Formwissen in den Segmentierungsprozess einzubeziehen, besteht bei der Verwendung von Radialstrahlen. Ausgehend von einem vorgegebenen Saatpunkt  $s$  werden  $n$  Strahlen radial gleichverteilt in alle möglichen Richtungen  $\mathbf{d}_i, i \in [1, \dots, n]$  ausgesendet. Eine Segmentierung kann nun als Oberfläche um den Saatpunkt beschrieben werden (siehe Abbildung 3.9), welche durch einen  $n$ -dimensionalen Radien-Vektor  $\mathbf{r}$  gegeben ist. Die Oberflächenpunkte  $\mathbf{p}_i$  sind durch  $\mathbf{p}_i = s + r_i \cdot \mathbf{d}_i$  gegeben. Diese Art der Repräsentation ist besonders für in etwa runde Strukturen, wie es z.B. Lymphknoten sind, geeignet. Da jeder Radialstrahl die Oberfläche nur einmal schneiden darf, ist diese Repräsentation konzeptbedingt allerdings auf die Beschreibung sternkonvexer Strukturen limitiert. Wird der Vektor  $\mathbf{r}$  diskretisiert ( $\mathbf{r} = \frac{r_{\max}}{m} \cdot \mathbf{l}, \mathbf{l} \in [1, \dots, m]$ ), kann der Segmentierungsprozess als Energieminimierungsprozess innerhalb eines *Markov-Random-Fields* (MRF) beschrieben werden<sup>1</sup>. Es soll die Belegung  $\mathbf{l}^*$  gefunden werden, die ein Energiefunktional der Form

$$E(\mathbf{l}) = \sum_{\forall i} \left( E(l_i) + \sum_{\forall j \in N(i)} E(l_i, l_j) \right) \quad (3.4)$$

minimiert, wobei  $N(i)$  die Menge aller benachbarter Radialstrahlen des Strahls  $i$  darstellt. Der Term  $E(l_i)$  hängt nur von der Belegung des  $i$ -ten Feldes (Strahls) ab und kann z.B. von dem Bildinhalt an der Position  $\mathbf{x} = s + l_i \frac{r_{\max}}{m} \mathbf{d}_i$  abhängen. Der Term  $E(l_i, l_j)$  hingegen ermöglicht es, lokales Formwissen in den Segmentierungsprozess zu integrieren, indem z.B. hohe Energiewerte für große absolute

---

<sup>1</sup> $r_{\max}$  bestimmt den maximal möglichen Radius.  $m \in \mathbb{N}$  bestimmt die Anzahl der möglichen Belegungen. Der Abstand zwischen zwei möglichen Radien ist folglich durch  $\frac{r_{\max}}{m}$  gegeben.

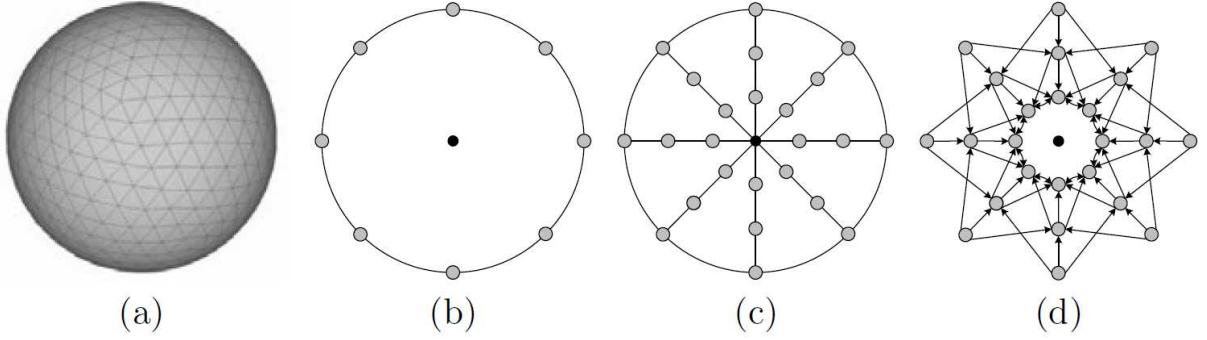


Abbildung 3.10.: Ausgehend von einem kugelförmigen Mesh (a) wird in mehreren Schritten der Graph erstellt (hier  $\Delta = 1$ , Abbildung aus [93]).

Differenzen benachbarter Belegungen  $l_i$  und  $l_j$  vergeben werden. Diese Art der Oberflächenrepräsentation wurde z.B. zur Segmentierung von neuroepithelialen Hirntumoren aus MRT-Aufnahmen eingesetzt [91], kommt aber auch zur Segmentierung von Lymphknoten zum Einsatz.

Barbu et al. verwenden in [92] und [67] solch eine Repräsentation zur Segmentierung von Lymphknoten aus CT-Aufnahmen. Zunächst wird für jeden Strahl  $i$  ein Radius  $y_i$  bestimmt, der aufgrund der Grauwertcharakteristik entlang des Strahls wahrscheinlich der Grenze des Lymphknotens entspricht:  $y_i = \underset{0 \leq r \leq R_{\max}}{\operatorname{argmin}} |I(\mathbf{s}) - I(\mathbf{s} + r\mathbf{d}_i)| > D_{\max}$ .  $y_i$  entspricht also der kürzesten Entfernung zum Zentrum, an dessen Position die Bildintensität von der des Zentrums um mehr als dem Schwellwert  $D_{\max}$  abweicht. Das Energiefunktional aus Gleichung 3.4 ergibt sich dann zu

$$E(\mathbf{r}) = \sum_{\forall i} \left( \alpha \rho(r_i - y_i) + \sum_{\forall j \in N(i)} (r_i - r_j)^2 \right), \quad (3.5)$$

wobei  $\alpha$  ein Gewichtungsfaktor und  $\rho(x) = \ln(1 + \frac{x^2}{2})$  eine nichtlineare Transferfunktion ist. Demnach werden Oberflächen bevorzugt, die sich nahe an den durch  $\mathbf{y}$  gegebenen Punkten befinden und gleichzeitig ähnliche benachbarte Radien aufweisen, also glatt sind. Die Optimierung erfolgt durch ein Gradientenabstiegsverfahren, welches durch  $\mathbf{r} = \mathbf{y}$  initialisiert wird. Der Evaluierung erfolgte auf einer sehr großen Anzahl von Lymphknoten (763), die sich allerdings nicht im Kopf-Hals-Bereich befanden. In Bezug auf eine manuelle Segmentierung wird ein DSC von  $0.78 \pm 0.13$  und eine durchschnittliche Oberflächendistanz von  $0.93 \pm 0.74 \text{ mm}$  erreicht.

Auch Wang et al. verwendeten in [93] Radialstrahlen zur Segmentierung von Lymphknoten aus CT-Aufnahmen. Dazu wird das in [94] vorgestellte allgemeine Verfahren zur optimalen Oberflächensegmentierung auf Radialstrahlen angewendet. Zunächst werden die Bildintensitäten des Eingabebildes in einem Vorverarbeitungsschritt so transformiert, dass nur noch Strukturen zwischen  $-100$  und  $150\text{-HU}$  – ein Bereich in dem sich Lymphknoten üblicherweise befinden – erhalten bleiben. Eine  $3 \times 3$ -Medianfilterung verringert Rauschen. Anschließend wird aus der diskreten Radialstrahlenrepräsentation ein gerichteter Graph erstellt, wie in Abbildung 3.10 dargestellt. Der Raum aller zulässigen Formen ist durch den Parameter  $\Delta$  eingeschränkt, welcher bestimmt, mit dem wievielten Knoten  $k_j = k_i - \Delta$  des benachbarten Strahls  $j$  ein Knoten  $k_i$  des Strahls  $i$  verbunden wird. Für  $\Delta = 0$  wird der Formenraum bspw. auf Kugeln beschränkt. Je größer  $\Delta$  ist, desto stärker kann

der Formenraum variieren. Den Kanten innerhalb eines Strahls werden Kosten zugeordnet und die eigentliche Segmentierung erfolgt durch einen *Graph-Cut*(GC) zwischen den Punkten am äußersten Rand und dem Mittelpunkt, so dass der Schnitt zu minimalen kumulierten Kosten führt. Die Kosten  $c_i(k)$  des  $k$ -ten Knotens des Strahls  $i$  ergeben sich aus Bildkantenkosten  $c_{\text{Kante}_i}(k)$  und Grauwert-Homogenitätskosten  $c_{\text{gh}_i}(k)$  zu  $c_i(k) = (1 - w_{\text{Form}}(k))c_{\text{Kante}_i}(k) + \alpha c_{\text{gh}_i}(k)$  mit dem Gewichtungsfaktor  $\alpha$ . Die einzelnen Terme sind durch

$$c_{\text{gh}_i}(k) = \max_{\forall j \leq k} I(\mathbf{s} + j\Delta\mathbf{d}_i) - \min_{\forall j \leq k} I(\mathbf{s} + j\Delta\mathbf{d}_i), \quad (3.6)$$

$$c_{\text{Kante}_i}(k) = \max_{a \in M_i} p(k, a) \quad (3.7)$$

und

$$w_{\text{Form}}(k) = \frac{\sum_{\forall i} c_{\text{Kante}_i}(k)}{\max_{\forall j} \sum_{\forall i} c_{\text{Kante}_i}(j)} \quad (3.8)$$

gegeben, wobei  $I(\mathbf{x})$  der Bildintensität in dem Punkt  $\mathbf{x}$  und  $M_i$  der Menge aller lokalen Maxima der Ableitung von  $c_{\text{gh}_i}(k)$  entspricht.  $p(k, a) = e^{-\frac{(k-a)^2}{2\sigma^2}}$  modelliert eine örtliche Unsicherheit.  $w_{\text{Form}}(k)$  ist besonders für solche Radien  $k$  hoch, für die die Kantenkosten aller Strahlen konsistent hohe Werte aufweisen. Benachbarte Strukturen, die den Saatpunkt nicht komplett umschließen, sorgen demnach nicht für ein hohes Gewicht. Die Verwendung von lokalen Maxima der Ableitungen der Bildhomogenität anstatt der sonst oft üblichen Bildgradientenstärke zur Beschreibung der Bildkantenkosten hat den Vorteil, dass starke Bildgradienten umliegender Strukturen nicht von den oft schwachen Bildgradienten an der Lymphknotengrenze ablenken können. Die Evaluierung erfolgte durch Vergleich des Segmentierungsergebnisses von 22 Lymphknoten mit einer manuell erstellten Referenzsegmentierung. Dabei wurde ein mittlerer DSC von 0.84, eine durchschnittliche symmetrische Oberflächendistanz von  $0.47 \pm 0.08 \text{ mm}$  und eine Hausdorff-Distanz von 2.12 mm erzielt.

Später wurde dieses Verfahren um die Möglichkeit einer interaktiven Korrektur erweitert [95]. Der Anwender klickt dazu nachdem ein erstes Segmentierungsergebnis, welches durch das gerade beschriebene Verfahren generiert wurde, auf einzelne Stellen der tatsächlichen Lymphknotenkontur, die nicht korrekt segmentiert wurden. Auf Grundlage dieses Punktes werden dann die bildbasierten Kosten in einer Region um diese Punkte, die sich durch Grauwertprofilähnlichkeiten definiert, angepasst. Der Strahl, welcher durch den Punkt verläuft, erhält minimale Kosten an dem Punkt selbst. Bei allen übrigen Strahlen der Region wird das Formgewicht  $w_{\text{Form}}$  so angepasst, dass ein ähnlicher Radius bevorzugt wird. Auf Grundlage dieser angepassten Kosten wird dann das Segmentierungsverfahren erneut aufgerufen. Die Evaluation wurde auf 111 Lymphknoten durchgeführt, eine deutliche Erweiterung im Vergleich zu [96]. Neben der Segmentierungsgenauigkeit wurde dabei auch die Abhängigkeit vom Saatpunkt zumindest rudimentär anhand einiger Beispiele betrachtet.

#### 3.1.6. Zusammenfassung

Sowohl die Segmentierung der Lymphknotenregionen als auch die Segmentierung einzelner Lymphknoten wird als Lymphknotensegmentierung bezeichnet. Der Fokus dieser Arbeit liegt auf Letzterem. Die meisten in der Literatur vorgestellten Verfahren werden auf CT-Daten angewendet, es existieren jedoch auch einige wenige Arbeiten für andere Bildmodalitäten, z.B. für MRT und Ultraschall. Die größten Herausforderungen bei der Segmentierung von Lymphknoten sind die hohe Varianz an Größe,

Form und an umliegendem Gewebe sowie dem oftmals geringen Kontrast zu benachbarten Strukturen und das Vorhandensein von Pathologien.

Da die ungefähre Objektgröße a priori nicht bekannt ist, haben Verfahren, die ein Modell an einen Datensatz mit Hilfe einer lokalen Strategie anpassen, einen großen Nachteil. Ist die initiale Modellgröße zu klein gewählt, kann die Objektgrenze nicht problemlos gefunden werden, weil die Gewebedichte in gutartigen Lymphknoten recht homogen ist und so z.B. eine Anziehung des Modells durch Bildkanten nicht möglich ist. Ist die initiale Modellgröße jedoch zu groß gewählt, ist eine fälschliche Anpassung an benachbarte Objekte nicht unwahrscheinlich. Häufig wird dieses Problem durch die Erhöhung der Anwenderinteraktion gelöst, es gibt jedoch auch andere Lösungsmöglichkeiten, wie z.B. die Verwendung mehrerer Modelle unterschiedlicher Größen. Die danach folgende automatische Modellauswahl ist allerdings nicht trivial und potentiell fehleranfällig. Darüber hinaus besteht die Möglichkeit der Verwendung globaler statt lokaler Optimierungsalgorithmen, wie z.B. Graph-Cut, die keine Initialisierung benötigen.

Als Bildmerkmale können Bildgradienten an der Objektgrenze bzw. Bildintensitäten im Inneren des Lymphknotens verwendet werden. Da das umliegende Gewebe sehr unterschiedlich sein kann, ist es nicht sinnvoll Annahmen darüber in den Segmentierungsprozess einzubeziehen. Dadurch entfällt z.B. auch das Lernen von Grauwertprofilen an den Objektgrenzen. Bei der Verwendung von Bildgradienten ist darauf zu achten, dass die Objektgrenzen umliegender Strukturen deutlich stärkere Bildgradienten als Lymphknoten aufweisen können. Bei gesunden Lymphknoten kann der Grauwertbereich im Inneren in einem bestimmten Intervall angenommen werden. Im Bereich der Objektgrenzen wird durch Partialvolumeneffekt allerdings davon abgewichen. Pathologische Lymphknoten enthalten hingegen oftmals nekrotisches Gewebe im Zentrum, welches sich deutlich dunkler als das typische Lymphknotengewebe im CT darstellt. Da gerade die Segmentierung pathologischer Lymphknoten von besonderer klinischer Bedeutung ist, sollte dies von einem Verfahren explizit berücksichtigt werden. Dies geschieht jedoch bei nur sehr wenigen in der Literatur vorgestellten Verfahren.

Da viele der umliegenden Strukturen (Muskeln, Gefäße, Drüsen, andere Lymphknoten, . . . ) sehr ähnliche Dichten wie Lymphknoten und damit auch sehr ähnliche Intensitäten in CT-Aufnahmen haben, ist der Kontrast am Übergang von Lymphknoten zu solchen Strukturen lokal nur sehr gering. Eine schichtweise 2D-Segmentierung auf axialen Schichten ist deshalb nicht zielführend. Daher werden echte 3D-Verfahren benötigt, die Formwissen einbeziehen.

Obwohl die Lymphknoten in etwa elliptisch sind, ist die Formenvarianz beträchtlich. Da jedoch kein bestimmter spezieller Formenraum bevorzugt oder ausgeschlossen werden kann, ist das Ausnutzen von Formkorrelationen unterschiedlicher Lymphknoten nicht möglich. Das Lernen des Formenraums bspw. durch die Hauptkomponentenanalyse (PCA) ist also nicht erfolgversprechend. Daher wird globales Formwissen in existierenden Ansätzen nur in Form von einfachen geometrischen Formen, wie z.B. Kugeln oder Ellipsoiden, eingebracht. Um dennoch alle möglichen Formen modellieren zu können, muss es allerdings möglich sein, stark von dieser Form abweichen zu können, was wiederum die Anfälligkeit zur fälschlichen Anpassung an umliegendes Gewebe erhöht. Alternativ wird lokales Formwissen verwendet, welches auch bei starker Gewichtung keine relevanten Formen ausschließt.

Da, wie in Tabelle 3.1 zu sehen ist, kein bestehendes Verfahren allen angesprochenen Aspekten genügt, besteht Bedarf an der Entwicklung eines neuartigen echten 3D-Verfahrens, welches Lymphknoten aus CT-Datensätzen mit minimaler Anwenderinteraktion (ein Klick) segmentiert. Um eine große Varianz an Lymphknoten abdecken und dem geringen Kontrast auf CT-Aufnahmen entgegentreten zu können, soll Wissen über das Aussehen im gesamten Inneren des Lymphknotens, insbesondere

| Verfahren         | Jahr | CT | Echtes 3D Verfahren | Segmentierung einzelner Lymphknoten | Ein-Klick-Initialisierung möglich | Verwendung von Wissen über das Aussehen | Verwendung von Formwissen | Globale Segmentierungsstrategie | Berücksichtigung von pathologischen Fällen | Anwendung auf klinische Daten | Vergleich mit Expertensegmentierung |
|-------------------|------|----|---------------------|-------------------------------------|-----------------------------------|---|---------------------------|---------------------------------|--|-------------------------------|-------------------------------------|
| Rogowska (AS)[74] | 1996 | ✓  | ✗                   | ✓                                   | ✓                                 | ✓                                       | ✗                         | ✓                               | ✗  | ✓                             | ✓                                   |
| Rogowska (LS)[74] | 1996 | ✓  | ✗                   | ✓                                   | ✗                                 | ✓                                       | ✗                         | ✗                               | ✗  | ✓                             | ✓                                   |
| Rogowska (WT)[74] | 1996 | ✓  | ✗                   | ✓                                   | ✗                                 | ✓                                       | ✗                         | ✗                               | ✗  | ✓                             | ✓                                   |
| Rogowska (SN)[74] | 1996 | ✓  | ✗                   | ✓                                   | ✗                                 | ✓                                       | ✓                         | ✗                               | ✗  | ✓                             | ✓                                   |
| Honea[87]         | 1997 | ✓  | ✗                   | ✓                                   | ✗                                 | ✗                                       | ✓                         | ✗                               | ✗  | ✓                             | ✗                                   |
| Honea[88]         | 1999 | ✓  | ✓                   | ✓                                   | ✓                                 | ✗                                       | ✓                         | ✗                               | ✗  | ✗                             | ✗                                   |
| Yan[89]           | 2004 | ✓  | ✗                   | ✓                                   | ✓                                 | ✓                                       | ✓                         | ✗                               | ✗  | ✓                             | ✗                                   |
| Dornheim[81]      | 2006 | ✓  | ✓                   | ✓                                   | ✓                                 | ✓                                       | ✓                         | ✗                               | ✓  | ✓                             | ✓                                   |
| Unal[76]          | 2006 | ✗  | ✗                   | ✓                                   | ✗                                 | ✓                                       | ✓                         | ✗                               | ✗  | ✓                             | ✓                                   |
| Kitasaka[64]      | 2007 | ✓  | ✓                   | ✓                                   | ✓                                 | ✓                                       | (✓)                       | ✗                               | ✗  | ✓                             | ✗                                   |
| Commowick[68]     | 2008 | ✓  | ✓                   | ✗                                   | -                                 | -                                       | -                         | -                               | -  | ✓                             | ✓                                   |
| Maleike[79]       | 2008 | ✓  | ✓                   | ✓                                   | ✗                                 | ✓                                       | ✓                         | ✗                               | ✗  | ✓                             | ✓                                   |
| Gorthi[70]        | 2009 | ✓  | ✓                   | ✗                                   | -                                 | -                                       | -                         | -                               | -  | ✓                             | ✓                                   |
| Moltz[78]         | 2009 | ✓  | ✓                   | ✓                                   | ✗                                 | ✓                                       | ✓                         | ✓                               | ✗  | ✓                             | ✗                                   |
| Zhang[77]         | 2009 | ✗  | ✗                   | ✓                                   | ✗                                 | ✓                                       | ✓                         | ✓                               | ✗  | ✓                             | ✓                                   |
| Dornheim[83]      | 2010 | ✓  | ✓                   | ✓                                   | ✓                                 | ✓                                       | ✓                         | ✗                               | ✓  | ✓                             | ✓                                   |
| Teng[71]          | 2010 | ✓  | ✓                   | ✗                                   | -                                 | -                                       | -                         | -                               | -  | ✓                             | ✓                                   |
| Wang[93]          | 2010 | ✓  | ✓                   | ✓                                   | ✓                                 | ✓                                       | ✓                         | ✓                               | ✗  | ✓                             | ✓                                   |
| Debats (CCRG)[75] | 2011 | ✗  | ✓                   | ✓                                   | ✓                                 | ✓                                       | ✗                         | ✗                               | ✗  | ✓                             | ✓                                   |
| Debats (ECC)[75]  | 2011 | ✗  | ✓                   | ✓                                   | ✓                                 | ✓                                       | ✗                         | ✗                               | ✗  | ✓                             | ✓                                   |
| Debats (GCS)[75]  | 2011 | ✗  | ✓                   | ✓                                   | ✓                                 | ✓                                       | ✓                         | ✓                               | ✗  | ✓                             | ✓                                   |
| Debats (PSAM)[75] | 2011 | ✗  | ✓                   | ✓                                   | ✓                                 | ✓                                       | ✓                         | ✗                               | ✗  | ✓                             | ✓                                   |
| Lu[85]            | 2011 | ✓  | ✗                   | ✓                                   | ✗                                 | ✗                                       | ✓                         | ✗                               | ✗  | ✓                             | ✓                                   |
| Barbu[67]         | 2012 | ✓  | ✓                   | ✓                                   | ✓                                 | ✓                                       | ✓                         | ✓                               | ✗  | ✓                             | ✓                                   |
| Beichel[95]       | 2012 | ✓  | ✓                   | ✓                                   | ✓                                 | ✓                                       | ✓                         | ✓                               | ✗  | ✓                             | ✓                                   |
| Feulner[90]       | 2012 | ✓  | ✓                   | ✓                                   | ✓                                 | ✓                                       | ✓                         | ✓                               | ✗  | ✓                             | ✗                                   |

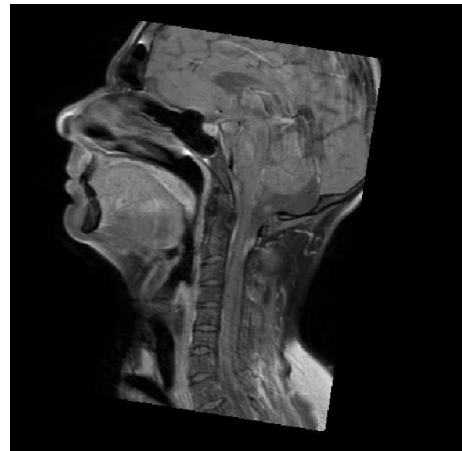
Tabelle 3.1.: Zusammenfassung von Eigenschaften der in diesem Kapitel betrachteten Segmentierungsverfahren von Lymphknoten



auch von pathologischen Lymphknoten, zur Anwendung kommen. Darüber hinaus soll Wissen über die Form durch lokales Formwissen einfließen und die Berechnung der Segmentierung soll durch ein globales Verfahren erfolgen.



(a) CT-Aufnahme in überstreckter Stellung



(b) MRT-Aufnahme in aufrechter Stellung

Abbildung 3.11.: Kopf-Hals-Aufnahmen des selben Patienten mit extrem unterschiedlicher Stellung des Nackens

## 3.2. Registrierung des Kopf-Hals-Bereichs

Die kritischste der in Kapitel 1.3 formulierten Zielsetzungen ist, dass zur Registrierung von Lymphknoten im Kopf-Hals-Bereich keine Annahme über die Stellung der Wirbelsäule getroffen werden kann. Dies hat zur Folge, dass derselbe Lymphknoten im Halsbereich bis zu 10 cm in beiden Aufnahmen entfernt sein kann (siehe Abbildung 3.11). Daher ist die bloße Anwendung von lokal beschränkten Registrierungsverfahren nicht möglich, sondern der gesamte Kopf-Hals-Bereich muss deformierbar registriert werden. Da lokale Verfahren im Anschluss angewendet werden können, spielt die Genauigkeit nur eine untergeordnete Rolle. Es kommt vielmehr auf die Robustheit gegenüber Bildartefakten und abgeschnittenen Aufnahmen, wie sie im klinischen Alltag entstehen, an.

Die Zielsetzung dieser Arbeit führt weiterhin dazu, dass für die Kopf-Hals-Registrierung innerhalb des in Kapitel 2.2.2 kurz vorgestellten Klassifikationsschemas nach Maintz et al. [52] folgende Festlegungen bereits an dieser Stelle getroffen werden können:

|                          |  |
|--------------------------|--|
| Dimensionalität:         | 3D/3D,   |
| Registrierungsgrundlage: | Grauwert- bzw. Segmentierungsbasiert,                  |
| Transformationsdomäne:   | Global,  |
| Interaktion:             | Keine (vollautomatisches Verfahren),                   |
| Beteiligte Modalitäten:  | CT/MRT bzw. CT/CT,                                     |
| Subjekt:                 | Intra-Subjekt, Anatomieänderungen möglich (Zeitreihe), |
| Objekt:                  | Lymphknoten bzw. Kopf-Hals-Bereich.                    |

Zu untersuchen bleibt also die Art der Transformation und das Optimierungsverfahren, wobei die Art der Transformation einen weit größeren Einfluss hat. Im folgenden Abschnitt werden in der Literatur beschriebene Verfahren zur Kopf-Hals-Registrierung beschrieben. Um dem Leser einen besseren Gesamteindruck zu vermitteln, wird dabei allerdings, wenn nötig, von der durch die obige Tabelle strikt vorgegebenen Eingrenzung abgewichen.

### 3.2.1. Transformationsklassen

**Rigide Verfahren:** Aufgrund der Starrheit des Kopfes ist dies das typische Anwendungsgebiet rigider Registrierungsverfahren. 1994 schlugen Hill et al. [97] eine landmarkenbasierte Methode zur Registrierung von CT- und MRT-Aufnahmen vor, für welche bei der Bildakquisition keine besonderen Vorkehrungen (wie z.B. ein Fixierahmen oder externe Landmarken) getroffen werden müssen. Zunächst identifiziert der Anwender mit Hilfe einer komfortablen graphischen Benutzeroberfläche korrespondierende 3D-Punkte in beiden Aufnahmen. Die Verschiebung wird durch die Schwerpunkte der Punktwolken und die Drehung durch eine Singulärwertzerlegung bestimmt. Das Verfahren wurde auf 15 CT/MRT-Bildpaaren getestet.

In den folgenden Jahren wurde von Studholme et al. in [98] und [99] ein vollautomatisches rigides CT/MRT-Registrierungsverfahren für den Kopf vorgestellt, welches ein Bildähnlichkeitsmaß in einem Multi-Resolution-Gradientenabstiegsverfahren maximiert. Dazu wurden die damals bekannten zur multimodalen Registrierung geeigneten Bildähnlichkeitsmaße *Soft Tissue Correlation* [100], *Corresponding MR and CT Variance* [101] sowie der *Mutual Information* [72][73] verwendet. Anhand von 10 Bildpaaren stellte sich heraus, dass zur CT/MRT Registrierung *Mutual Information* und *Soft Tissue Correlation* am besten geeignet sind und vergleichbare Genauigkeit wie landmarkenbasierte Verfahren liefern. Auch die Registrierung von abgeschnittenen Volumen ist bis zu einem gewissen Grad möglich. Das Registrierungsergebnis lässt sich verbessern, indem deformierbare Teile, wie z.B. das Ohr oder der Unterkiefer, nicht zur Berechnung des Bildähnlichkeitsmaßes herangezogen werden.

Mittlerweile hat sich diese Art der Registrierung zu einem Standardverfahren entwickelt, welches in vielen kommerziell verfügbaren Systemen zum Einsatz kommt. Bei der Wahl des Fehlermaßes wird *Mutual Information* nach wie vor häufig verwendet, obwohl andere vielversprechende Verfahren wie *Normalized Mutual Information* [102], *Gradient Correlation* [103], *Cumulative Residual Entropy* [104] und *Cross-Cumulative Residual Entropy* [105] in der Zwischenzeit publiziert wurden.

Wie in Kapitel 2.2.2.1 erörtert, besitzen rigide Verfahren eine geringe Komplexität und sind durch die Verwendung von globalem Vorwissen sehr robust gegenüber lokalen Bildstörungen und deformieren das Bild nicht. Allerdings ist anhand des Beispiels in Abbildung 3.11 leicht zu sehen, dass diese Transformationsklasse für die Registrierung des gesamten Kopf-Hals-Bereichs – und insbesondere der Lymphknotenregionen im Hals – aufgrund von möglichen Deformationen ungeeignet ist. Sollen dennoch rigide Registrierungsverfahren eingesetzt werden, muss sichergestellt werden, dass die Position des Halses in allen Aufnahmen dieselbe ist. Dies geschieht in der klinischen Praxis üblicherweise durch einen stereotaktischen Rahmen, der den Patienten bei allen Aufnahmen in derselben Position fixiert.

**Spline-basierte und freie elastische Verfahren:** Im Kopf-Hals-Bereich werden Spline-basierte und komplett freie elastische Verfahren hauptsächlich zur Inter-Subjekt-Registrierung von CT-Aufnahmen im Kontext der atlasbasierten Registrierung verwendet. Haupteinsatzgebiet ist dabei die automatische Konturierung von Risikoorganen, wie z.B. der Speicheldrüsen, und Zielorganen, wie z.B. die Lymphknotenregionen, zur Strahlentherapieplanung. Im Folgenden wird eine Auswahl an Verfahren kurz vorgestellt, die in den letzten Jahren publiziert wurden. Um nicht abzuschweifen, wird dabei lediglich auf die Registrierungskomponente der Verfahren eingegangen und andere Komponenten, wie die Auswahl des Atlas und der Segmentierungsschritt, werden ausgeklammert.

Hollensen et al. [106] registrieren zunächst affin und finden anschließend ein Minimum des qua-

dratischen Intensitätsunterschieds einer B-Spline-Transformation unter Verwendung von lokaler Regularisierung. Bei Ramus et al. [107] kommt ebenfalls eine affine Registrierung zum Einsatz. Im Anschluss daran wird der in [108] vorgestellte Multi-Resolution-Block-Matching-Algorithmus verwendet, welcher ein dichtes Deformationsfeld aus dem Versatz (also nur Translation) einzelner Blöcke berechnet. Ausreißer einzelner Blocks werden detektiert und durch Interpolation der Transformationen gültiger Blocks ersetzt. Gering et al. [109] verwenden die klassische Demons-Registrierung [58]. Nach einer rigiden Registrierung des Halswirbels C2, verwenden Yang et al. [110] eine Art Multi-Resolution-Demons-Ansatz, der in Bezug auf Genauigkeit und Geschwindigkeit, insbesondere in Regionen mit schwachem Bildgradienten, verbessert wurde. Papastavrou et al. [111] registrieren PET-CT-Aufnahmen verschiedener Patienten in einem Framework, das zunächst die Bilder affin registriert und anschließend die Kontrollpunkte einer B-Spline-Transformation mit einem Gradientenabstiegsverfahren optimiert. Als Bildähnlichkeitsmaß wird eine gewichtete Summe aus der quadratischen Intensitätsdifferenz der CT-Aufnahmen und der normalisierten Mutual Information zwischen der restlichen Kombination aus CT- und PET-Aufnahmen verwendet. Es kommt ebenfalls lokales Vorwissen zum Einsatz.

Darüber hinaus gibt es Ansätze, die neben der Intensitätsinformation der Bilder auch die bereits konturierten Strukturen des Atlas einbeziehen. Commowick et al. [68] registrieren anfangs affin und anschließend die im Atlas vorhandenen Strukturen lokal-affin. Die Deformationen in allen anderen Bereichen werden zunächst durch Interpolation bestimmt. Zur Bestimmung der endgültigen Transformation für jeden einzelnen Voxel wird ein zu [108] ähnlicher Multi-Resolution-Block-Matching-Algorithmus verwendet. Han et al. [112] registrieren in 3 Stufen: 1. Maximierung der Mutual Information mit 9 Freiheitsgraden (Verschiebung, Drehung, Skalierung), 2. Blockbasierte Registrierung die die Glattheit der Objektoberflächen forciert (*Poly-Smooth*) und 3. Komplette freie elastische Registrierung, welche die Summe der *Local Correlation Coefficients* [113] mit Hilfe eines Gradientenabstiegsverfahrens minimiert. Eine lokale Regularisierung hat innerhalb der Objekte einen deutlich höheren Einfluß als in der Umgebung. Gorthi et al. [70] kombinieren eine komplett freie Registrierung mit aktiven Konturmodellen der Atlas-Strukturen, welche implizit durch Level-Sets beschrieben werden. Eine affine Vorregistrierung wird hierbei nicht benötigt. Teng et al. [71] verwenden globale Regularisierung, indem sie automatisch segmentierte Landmarken in einen Multi-Resolution B-Spline basierten Ansatz [114] einbeziehen, der die Mutual Information maximiert.

Aufgrund der hohen Variabilität der Anatomie in dem Inter-Subjekt-Fall müssen dabei sehr flexible Registrierungsverfahren zum Einsatz kommen, die sich, wie in Abschnitt 2.2.2.1 erörtert, durch eine hohe Komplexität und eine damit verbundene geringe Robustheit auszeichnen. Dieser hohe Grad an Flexibilität wird allerdings in dem in dieser Arbeit betrachteten Intra-Subjekt-Fall gar nicht benötigt. Daher ist zur Intra-Subjekt-Registrierung eine Transformationsklasse mit geringer Komplexität und höherer Robustheit vorzuziehen. Die Robustheit kann weiterhin durch die Verwendung von globalem anstatt lokalem Vorwissen über die Transformation gesteigert werden.

**Multi-Rigide Verfahren:** Multi-rigide Verfahren bieten für die Intra-Subjekt-Registrierung einen guten Kompromiss zwischen der Robustheit von rigider Registrierung auf der einen und der Flexibilität von elastischer Registrierung auf der anderen Seite, wodurch diese besonders geeignet für den in dieser Arbeit betrachteten Anwendungsfall erscheinen. Daher wird auf diese Transformationsklasse etwas umfangreicher eingegangen.

Alle multi-rigiden Verfahren, die im Kopf-Hals-Bereich Anwendung finden, betrachten ausschließ-

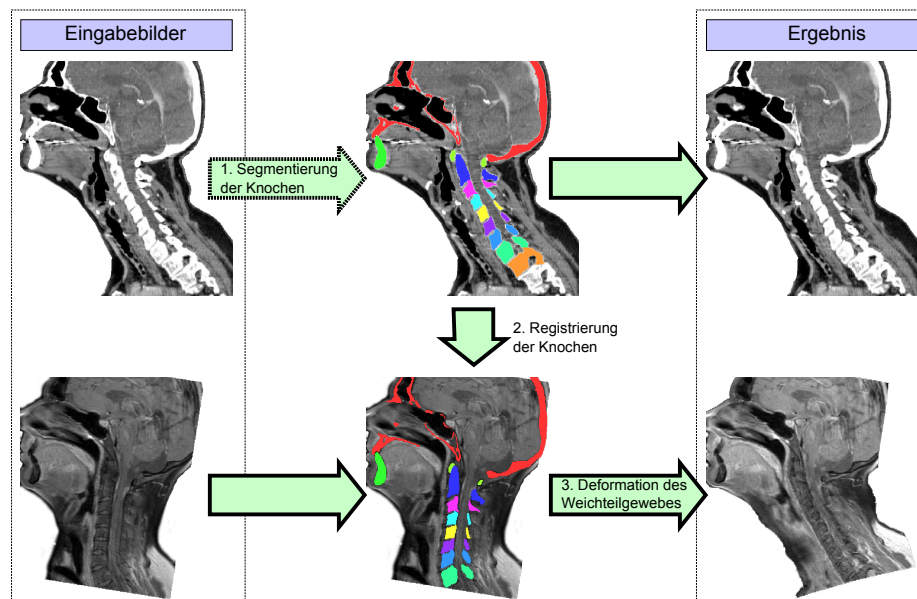


Abbildung 3.12.: Prinzipieller Ablauf einer multi-rigiden Registrierung

lich den Intra-Subjekt-Fall, wobei dies entweder multimodale Registrierung oder Registrierung einer Zeitserie bedeuten kann. Die einzelnen Knochen werden dabei als rigide Strukturen angesehen und das umliegende Weichteilgewebe als deformierbar. Im Intra-Subjekt-Fall ist diese Annahme i.d.R. zutreffend, da Knochen deutlich härter als das umliegende Gewebe und daher nicht deformierbar sind. Ferner ändert sich die Anatomie, bspw. durch Tumorwachstum, im Weichteilbereich deutlich stärker.

Die erste Arbeit in diesem Bereich wurde bereits im Jahre 1996 von Little et al. [54] vorgestellt. Auf der Basis von bekannten rigiden Transformationen einzelner Bereiche und bekannten korrespondierenden Landmarken wird ein Deformationsfeld durch Interpolation berechnet, welches die Starrheit der rigiden Teile aufrechterhält. Neben der Anwendung auf künstlichen 2D-Daten wurde dieses Verfahren auch zur Deformation von einzelnen sagittalen MRT-Schichten des Kopf-Hals-Bereichs verwendet. Eine Evaluation der Genauigkeit, insbesondere in den interpolierten Regionen, fand nicht statt.

Fast ein Jahrzehnt später, nämlich im Jahre 2005, stellten Du Bois d'Aische et al. in [115] eine gekoppelte Registrierung einzelner rigider Teile vor, welche in [116] und [117] weiterentwickelt wurden. Zunächst werden die einzelnen rigiden Strukturen von einem Anwender interaktiv segmentiert. Der Hauptbeitrag dieser Arbeit besteht darin, dass die einzelnen Strukturen dann nicht individuell, sondern gemeinsam registriert werden. Dies entspricht einer lokalen Regularisierung und erhöht den Konvergenzbereich der Methode erheblich. Die Deformation der nicht-rigiden Bildbereiche wird durch eine biomechanische Finite-Elemente-Methode (FEM) bestimmt. Die Methode wurde auf 7 Datensatzpaaren (4 CT/MRT, 2 CT/PET, 1 CT/CT) evaluiert, wobei bei den CT/MRT Paaren der Target-Registration-Error (TRE) von jeweils 20 anatomischen Landmarken bestimmt wurde. Gegenüber rigider Registrierung wurde über eine durchschnittliche Verbesserung der Genauigkeit um 48% auf 2.53 mm berichtet. Die Registrierung benötigte auf damals aktueller Hardware (AMD Athlon XP 3200+, 1 GB RAM) ca. 15 min.

Urschler et al. stellten im folgenden Jahr ein multi-rigides Verfahren [118] zur Registrierung und

Segmentierung von Knochen in Kopf-Hals-CT-Aufnahmen für die Angiografie. Zunächst wird die kontrastverstärkte mit der gewöhnlichen CT-Aufnahme rigide registriert. Der nicht korrekt registrierte Bereich wird dann über die durch ein Schwellwertverfahren erhaltenen Segmentierungen der Knochen bestimmt. Dieser Bereich wird erneut rigide registriert. Dieser Prozess wird solange fortgesetzt, bis der nicht korrekt registrierte Bereich eine vorgegebene Mindestgröße unterschreitet. Die Transformation des Gewebes wird dann durch Interpolation der einzelnen rigiden Transformationen bestimmt. Da die Segmentierung der Knochen in beiden Aufnahmen erforderlich ist, ist eine Anwendung auf den multi-modalen Fall nicht möglich. Die Methode wurde auf 5 Datensatzpaaren evaluiert, indem die Anzahl der nicht korrekt registrierten Knochenvoxel gezählt und die mittlere quadratische Abweichung berechnet wurde. Wie zu erwarten, verringerten sich beide Maße mit Zunahme der Iterationen. Die Laufzeit befindet sich im Bereich von wenigen Minuten (Pentium M 2.0 GHz, 1.5 GB RAM).

Im Jahre 2010 verwendeten Al-Mayah et al. ebenfalls ein multi-rigides Verfahren [119] zur Registrierung von Kopf-Hals-CT-Aufnahmen desselben Patienten zur Strahlentherapieplanung. Die einzelnen rigiden Strukturen (Halswirbel und Unterkiefer) werden manuell segmentiert und anschließend individuell rigide registriert. Zusätzlich kann auch die Patientenhautoberfläche weitere korrespondierende Landmarken liefern, die in den Registrierungsprozess einbezogen werden können. Das Deformationsfeld wird ebenfalls durch FEMs bestimmt. Die Methode wurde an vier Patienten evaluiert. Dabei wurden jeweils ein Tumor und die Ohrspeicheldrüsen in beiden Aufnahmen manuell segmentiert. Nach Anwendung der Transformation wurde der *Dice Similarity Coefficient* (DSC) und der Fehler des Massezentrums dieser Strukturen gemessen. Die Zunahme des DSCs gegenüber rigider Registrierung war mit knapp 15% auf 0.86 ohne die Hinzunahme der Körperoberfläche am größten. Bei dem Massezentrenfehler führte die Hinzunahme der Körperoberfläche allerdings zu einem besseren Ergebnis. Gegenüber rigider Registrierung wurde dieser um 38.8% auf 2.27 mm gesenkt.

Ein verwandtes Anwendungsgebiet der multi-rigiden Registrierung ist die Wirbelsäule, welches in [120] [121][122] und [123] angegangen wurde. Auch wurden multi-rigide Verfahren zur Registrierung von Ganzkörper-CT-Aufnahmen verwendet [124]. Bei Mäusen kam diese Art der Registrierung am ganzen Körper [125] bzw. an Extremitäten [126][127] zum Einsatz. Zu guter Letzt zählt die Registrierung auch Röntgenaufnahmen der Hand [128] und MRT-Aufnahmen des Gehirns [53] zu den Anwendungsgebieten. Da diese Verfahren nicht auf den Kopf-Hals-Bereich angewendet wurden, seien sie an dieser Stelle nur kurz erwähnt. Eine detaillierte Beschreibung relevanter Aspekte für diese Arbeit findet später statt.

Wie in Kapitel 2.2.2.1 bereits erwähnt, besteht eine automatisches multi-rigides Registrierungsverfahren aus drei Schritten, nämlich i) *Segmentierung der einzelnen rigiden Strukturen im Referenzbild*, ii) *Registrierung der rigiden Strukturen* und iii) *Bestimmung der Deformation in den restlichen Strukturen*. Deren Zusammenhang ist in Abbildung 3.12 dargestellt. In den folgenden Abschnitten werden nun konkrete in der Literatur vorgestellte Möglichkeiten dieser Schritte erörtert.

#### 3.2.2. Bestimmung der einzelnen Rigiden Strukturen

Eine der größten Herausforderungen bei diesem Ansatz ist die Identifikation der rigiden Teile  $R_i$ , also der einzelnen Knochen. Dabei ist es bei den meisten Verfahren nicht ausreichend für jedes Voxel des Referenzbildes zu bestimmen, ob es zu einer rigiden Struktur gehört, sondern auch zu welcher Struktur. Während Ersteres für Knochen auf CT-Bildern durch ein einfaches Schwellwertverfahren robust bestimmt werden kann, ist Letzteres ungleich komplizierter, weil einzelne rigide Strukturen,

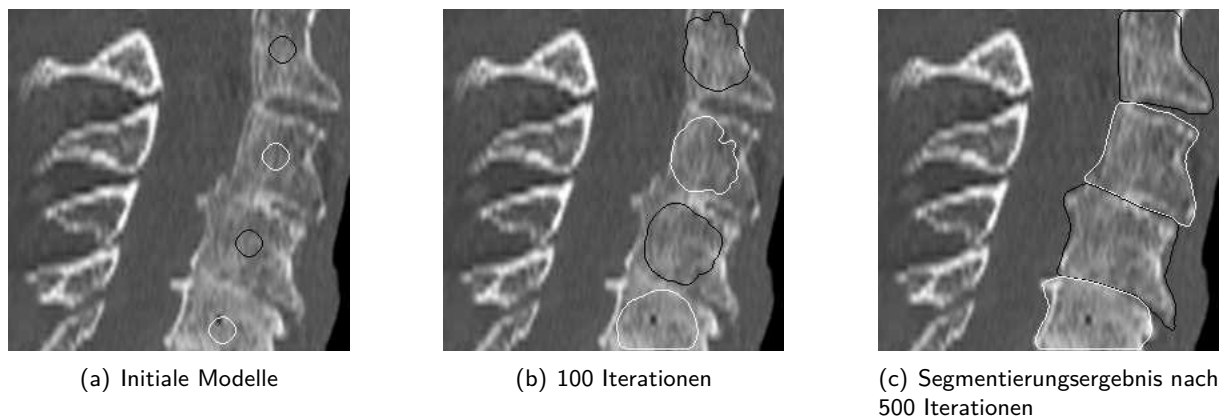


Abbildung 3.13.: Gleichzeitige Anpassung mehrerer Modelle an die Wirbelkörper zur Bestimmung der einzelnen rigiden Strukturen in einem multi-rigiden Registrierungsverfahren. Ein Startpunkt pro Wirbel ist erforderlich. (Abbildungen aus [121])

wie z.B. die Wirbel, sehr ähnlich aussehen und aneinander grenzen.

Ausnahmen hierzu bilden die Verfahren von Li et al. [124] und Suh et al. [127], welche keine explizite Zuordnung der rigiden Voxel zu einzelnen rigiden Teilen benötigen. Allerdings setzen beide Verfahren voraus, dass die rigiden Strukturen sowohl im Referenzbild als auch im Modellbild identifiziert werden können. Da dies nur in CT-Aufnahmen einfach zu bestimmen ist, eignen diese Verfahren sich nicht zur multimodalen Registrierung. In beiden Verfahren werden zunächst die Knochen in beiden Bildern mittels Schwellwertverfahren segmentiert. In [124] wird ein robuster punktbasierter Registrierungsalgorithmus [129] auf aus den Knochen extrahierten Punktwolken angewendet, welcher iterativ Punktkorrespondenz und die Registrierungstransformation bestimmt. In [127] wird ein erweiterter Demons-Algorithmus zur Registrierung der Knochenoberflächen verwendet. Obwohl diese Verfahren aufgrund der Verwendung von nicht-rigiden Registrierungsverfahren zur Registrierung der Knochen streng genommen nicht multi-rigide sind, werden diese dennoch als solche betrachtet, weil rigide Strukturen (die Knochen) und Weichteilgewebe unterschiedlich betrachtet werden.

**Manuelle / Interaktive Segmentierung:** Ein Großteil der bereits genannten multi-rigiden Registrierungsverfahren geht überhaupt nicht auf diesen Aspekt ein, sondern setzt die Kenntnis der einzelnen rigiden Strukturen als gegeben voraus. Dazu gehören [54][63][123][53] und [119].

Interaktive Verfahren kommen in [120], [126] und [117] zum Einsatz. In [120] werden einzelne Wirbel durch eine Kombination aus Schwellwertverfahren, Region-Growing und manueller Nachverarbeitung segmentiert. In [126] werden die einzelnen Knochen von Mäuseextremitäten durch die manuelle Platzierung von Trennebenen in den Gelenken – also den Übergängen der einzelnen rigiden Teile – voneinander getrennt. In [117] wird zur Segmentierung der Knochen aus Kopf-Hals-CT-Aufnahmen ein Softwaretool zur Simulation und Navigation in der Chirurgie [130] verwendet.

**Semi-automatische Verfahren:** Hu et al. stellten in [121] ein semi-automatisches Verfahren zur Segmentierung der einzelnen Wirbel der Halswirbelsäule aus CT-Aufnahmen vor. Zunächst setzt der Anwender je einen Startpunkt in jeden Wirbelkörper. Um diese Punkte werden kleine kugelför-

mige Modelle platziert, welche durch Level-Sets beschrieben werden. Die Modelle passen sich dann gleichzeitig an die Wirbelkörper an (siehe Abbildung 3.13). In die externe Energie fließen sowohl die Bildgradientenstärke als auch die Bildhelligkeit ein. Letzteres zum einen, um zu verhindern, dass die Kontur typische Knochenintensitäten verlässt und zum anderen, um die Anpassungsgeschwindigkeit im dunkleren inneren Bereich (Knochenbälkchen) zu erhöhen. Zusätzlich wird die Anpassung dahingehend verändert, dass die Anpassungsgeschwindigkeit in Bereichen, in welchen Konturen verschiedener Wirbelkörper sich nahe kommen, verringert wird, um Überschneidungen zu verhindern. Sich dennoch überschneidende Bereiche stoßen sich voneinander ab. Eine Evaluation fand lediglich qualitativ durch visuelle Begutachtung einiger weniger Datensätze statt.

**Automatische Heuristiken:** Die automatischen Verfahren von Čech et al. [122] zur Registrierung der Wirbelsäule und dem von Martín-Fernández et al. [128] zur Segmentierung von Händen auf Röntgenaufnahmen basieren auf komplizierten Heuristiken, wodurch sich diese Verfahren nur für genau den behandelten Anwendungsfall eignen. Die Verfahren sind also nicht generalisierbar.

In [122] kommt, nachdem sämtliche Knochen in der CT-Aufnahme durch ein einfaches Schwellwertverfahren segmentiert wurden, eine Folge von elementaren Bildverarbeitungsschritten zur Extraktion der Mittellinie des Rückenmarks zum Einsatz. Anschließend wird das Volumen in Schichten orthogonal zu der Knochenmarkmittellinie abgetastet, so dass ein normalisiertes Volumen entsteht, in dem die Wirbelsäule begradigt ist. Nach einer weiteren Folge von Verarbeitungsschritten zur robusteren Segmentierung der Knochen werden die einzelnen Wirbel durch Trennoberflächen voneinander separiert. Die Platzierung dieser Oberflächen erfolgt in Schichten mit einem lokalen Minimum des Knochenanteils, was den Bandscheiben entspricht. Die Evaluierung erfolgte nur durch visuelle Begutachtung der Ergebnisse von 10 Datensätzen. Die Autoren berichten von Problemen, insbesondere bei der Separierung der Halswirbel.

In [128] werden zur Registrierung von Händen auf Röntgenaufnahmen korrespondierende Landmarken in den Gelenken durch eine Heuristik extrahiert. Dies geschieht durch eine Folge von einfachen Bildverarbeitungsoperationen, darunter Gaußsche Filterung, Kantendetektion, Schwellwertverfahren und die Wasserscheidentransformation.

**Automatische generalisierbare Verfahren:** Ein generalisierbarer Ansatz zur Segmentierung von einzelnen rigiden Strukturen wird lediglich von Baiker et al. [125][131] zur Registrierung von Ganzkörper-CT-Aufnahmen von Mäusen angewendet. Auf einem repräsentativen hochaufgelösten CT-Datensatz wurden die einzelnen Knochen segmentiert. Dieser fungiert als artikulierter Atlas. Die mögliche Beweglichkeit wird durch eine hierarchische Modellierung des Skeletts erreicht, wobei auf verschiedenen Modellebenen Knochenverbunde bzw. einzelne Knochen den Modellelementen entsprechen. Zwischen den einzelnen rigiden Strukturen werden Gelenke explizit modelliert (siehe Abbildung 3.14(b)). Dabei wird zwischen Drehgelenken, Kugelgelenken und dem Schultergelenk in Bezug auf mögliche Freiheitsgrade unterschieden. Die Oberflächen der Knochen werden durch Punktwolken beschrieben. In der zu segmentierenden CT-Aufnahme werden die Knochen ebenfalls durch eine Punktwolke beschrieben. Der Anpassungsprozess des Atlas an die CT-Aufnahme erfolgt hierarchisch, wobei nach einer Ganzkörperausrichtung mit dem Kopf begonnen wird (siehe Abbildung 3.14(a)). Die Punktwolken der einzelnen Modellelemente werden mit der Knochenoberflächenpunktwolke sukzessive gemäß der spezifizierten Freiheitsgrade des Gelenks mittels *Iterative Closest Point* (ICP) [132] registriert. Lediglich die Wirbelsäule wird anderweitig segmentiert und zwar durch ein einfaches



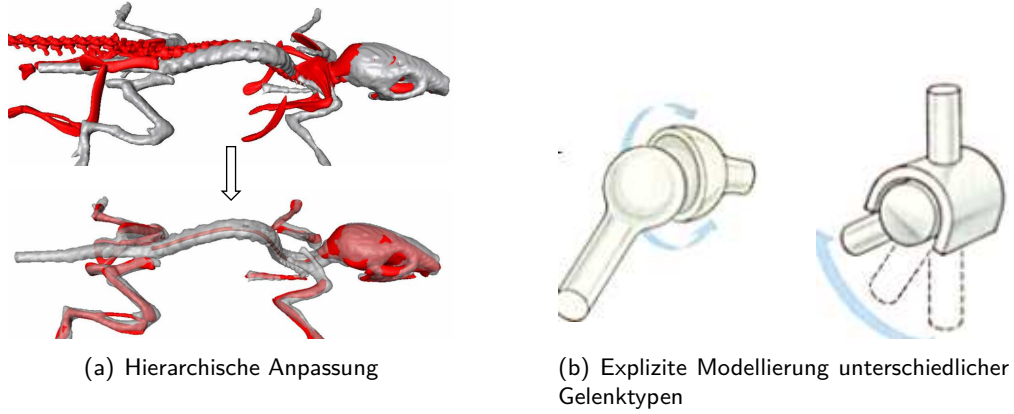


Abbildung 3.14.: Ein hierarchischer artikulierter Atlas wird von Baiker et al. [125] zur Segmentierung des Skeletts von Mäusen vorgestellt. (Abbildungen aus [125])

Region-Growing, wodurch die komplexe Modellierung der Wirbelsäule entfällt. Zur Evaluierung wurde der Ansatz auf 12 Ganzkörper-CT-Aufnahmen von Mäusen angewendet. Die durchschnittliche Oberflächendistanz der Knochen betrug nach der Anpassung 0.44 mm mit einer Standardabweichung von 0.4 mm. Das hierarchische Vorgehen führt zu einem großen Konvergenzbereich. Allerdings wird dieser dadurch erkauft, dass der Registrierungserfolg der unteren Hierarchieebenen von dem Erfolg in den oberen Ebenen abhängt.

Klinder et al. stellen einen Ansatz zur gemeinsamen Segmentierung von Wirbeln mittels eines artikulierten Formmodells vor [133]. Obwohl die beabsichtigte Anwendung der Autoren nicht die multi-rigide Registrierung ist, soll der Ansatz aufgrund der Nähe zu der für diese Arbeit benötigten Art der Segmentierung dennoch beschrieben werden. Das artikuliert Formmodell besteht aus je einem eigenen Modell für jeden Wirbel, welches in einem zweistufigen Prozess an einen CT-Datensatz angepasst wird: Eine globale Anpassung bestimmt Position und Ausrichtung der Wirbel und eine lokale Anpassung bestimmt anschließend gemeinsam die Form der Wirbel. In beiden Schritten wird der mögliche Lösungsraum gelernt. Für die lokale Anpassung geschieht dies mit Hilfe eines gewöhnlichen ASMs. Die globale Anpassung hingegen erfordert eine Beschreibung wahrscheinlicher Konstellationen der lokalen Koordinatensysteme der Wirbel. Da rigide Transformationen nicht zu einem Vektorraum gehören und bspw. eine Durchschnittstransformation nicht ohne Weiteres angegeben werden kann, kann dies allerdings nicht direkt mit Hilfe der in Kapitel 2.2.1 vorgestellten Methoden zur Segmentierung mit globalem Formwissen erfolgen. Daher sei zunächst das Distanzmaß zweier Transformationen  $T_1$  und  $T_2$  nach [134] definiert:

$$d(T_1, T_2) := \sqrt{\|\mathbf{r}(T_2^{-1} \circ T_1)\|^2 + \|0.05 \cdot \mathbf{t}(T_2^{-1} \circ T_1)\|^2}, \quad (3.9)$$

wobei  $\mathbf{r}(\cdot)$  der durch einen Einheitsquaternion beschriebenen Rotationskomponente und  $\mathbf{t}(\cdot)$  der Translationskomponente entsprechen. Der Mittelwert  $\mu$  von  $N$  Trainingstransformationen  $T^i$  wird mit Hilfe der Exponentialabbildung (vgl. Abschnitt 4.2.2) bzgl. der Distanz aus Gleichung 3.9 iterativ durch

$$\mu_{\mu^k}^{k+1} = \exp_{\mu^k} \left( \frac{1}{N} \sum_{i=0}^{N-1} \log(T^i) \right) \quad (3.10)$$

angenähert. Ein Vektor bestehend aus den individuellen Transformationen  $T_k$  relativ zu dem Nachbarwirbel beschreibt eine Wirbelsäulenkonstellation  $\mathbf{C} = [T_1, \dots, T_K]$ . Die durchschnittliche Konstellation  $\bar{\mathbf{C}}$  kann analog gebildet werden. In dem sich dort befindlichen Tangentialraum wird nun mittels PCA die Variation aller auftretenden Wirbelsäulenkonstellationen gelernt. Aus der dadurch erhaltenen Eigenvektormatrix  $A = (\mathbf{a}_1, \dots, \mathbf{a}_n)$  und dem aus einer bestimmten Konstellation  $\mathbf{C}$  bestimmten  $t$ -dimensionalen Gewichtsvektor  $\mathbf{b} = A^T(\log_{\bar{\mathbf{C}}} \mathbf{C})$  wird für die relative Transformation  $T_k$  des Elements  $k$  eine interne Energie  $E_{int_k}$  für den Modellanpassungsprozess abgeleitet:

$$E_{int_k} = d \left( \exp \left( \sum_{i=1}^t b_i \mathbf{a}_{i,k} \right), T_k \right)^2, \quad (3.11)$$

welche unwahrscheinliche Konstellationen, also Konstellationen mit einer großen Distanz zum gelernten Konstellationsraum, bestraft. Zur Anpassung an das Bild wird eine externe Energie  $E_{ext_k}$  eingeführt. Durch einen *Exhaustive Search* in einem Bereich um die aktuelle Transformation (sowohl bzgl. Verschiebung als auch bzgl. Drehung) wird in jedem Iterationsschritt für jedes Modellelement diejenige absolute Transformation  $\tilde{T}_{abs_k}$  gefunden, welche aufgrund der Bildgradientenstärke in der Modelloberfläche am besten zu dem Bild passt. Diese werden dann in relative Transformationen  $\tilde{T}_k$  bzgl. des Nachbarwirbels umgerechnet. Die so entstandenen *Merkmals Transformationen* dienen dann als Referenz zur Berechnung der externen Energie  $E_{ext_k}$  einer Transformation  $T_k$  des Elements  $k$ :

$$E_{ext_k} = d(\tilde{T}_k, T_k)^2. \quad (3.12)$$

Zur globalen Anpassung des artikulierten Modells an den Datensatz wird die gewichtete Summe aus externer und interner Energie mit Hilfe eines *Downhill-Simplex*-Optimierers minimiert. Die darauf folgende lokale Formanpassung geschieht bei allen Wirbeln gleichzeitig, wobei ähnlich zu [121] die externe Energie um einen Term erweitert wird, welcher Überlappung verhindert. Zur Evaluierung wurde die Methode erfolgreich auf 10 CT-Datensätze der Wirbelsäule angewendet, wobei die durchschnittliche Oberflächendistanz 1.0 mm betrug.

Durch das Einbringen von Objektlagewissen durch den gelernten Konstellationsraum können nicht plausible Konstellationen ausgeschlossen werden. Dies erhöht die Robustheit des Algorithmus im Vergleich zur unabhängigen Modellanpassung für jeden Wirbel. Die Wahrscheinlichkeit einer fälschlichen Anpassung eines einzelnen Wirbelmodells an einen ähnlichaussehenden Nachbarwirbel wird verringert. Folgende Teilaspekte des Algorithmus bieten allerdings Raum für Verbesserung:

- Als interne Energie wird die Distanz zum gelernten Konstellationsraum verwendet (DFFS, vgl. Abschnitt 4.2.3). Jedoch sind nicht alle Konstellationen innerhalb des Konstellationsraums plausibel. Je weiter eine Konstellation von der Durchschnittskonstellation entfernt ist, desto unplausibler ist sie. Dies (DIFS) wird nicht berücksichtigt.
- Für jedes Modellelement wird eine geeignete Transformation durch einen *Exhaustive Search* gefunden. Die globale gemeinsame Anpassung aller Elemente erfolgt dann durch einen *Downhill-Simplex*-Optimierer. Dies zieht vermutlich eine hohe Laufzeit nach sich. Eine geschlossene Ableitung des zu minimierenden Gesamtenergiefunktionalen würde den *Exhaustive Search* vermeiden und dabei die Verwendung gradientenbasierter Optimierungsverfahren ermöglichen. Eine Verringerung der Laufzeit wäre die Folge.
- Das gleichzeitige Finden einer Transformation für jeden Wirbel erfordert eine gute Initialisierung. Andernfalls kann es trotz Berücksichtigung des Konstellationsraums passieren, dass das

Modell teilweise um einen Wirbel *verrutscht*. Dies geschieht bspw. wenn das initiale Modell zu groß ist. Die oberen Wirbel könnten korrekt liegen, mehrere weiter unten liegende Wirbel jedoch nicht. Die Folge wäre, dass sich die unteren Modellelemente fälschlicherweise an die Nachbarwirbel anpassen könnten und es zu einer *Lücke* zwischen zwei Modellelementen in der Mitte des Gesamtmodells kommen könnte, die aufgrund der geringen externen Energie nicht durch die höhere interne Energie ausgeglichen werden könnte. Ein Start der Anpassung mit einem einzelnen bestimmten Modellelement und sukzessiver Einbeziehung von Nachbarmodellelementen könnte dies vermeiden.

### 3.2.3. Registrierung der rigiden Strukturen

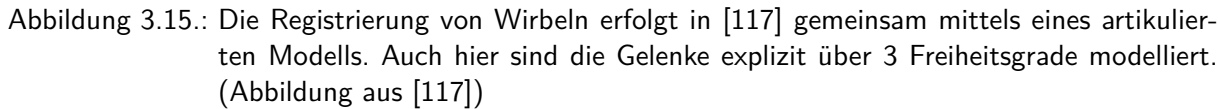
Durch die Annahme, dass Knochen sich nicht deformiert, ist es in dem Intra-Subjekt-Fall ausreichend, dass für jeden zuvor segmentierten Knochen  $i$  genau eine rigide Transformation  $T_i$  gefunden wird. Bei der Registrierung zweier CT-Aufnahmen ist dazu eine Registrierung der einzelnen rigiden Strukturen nicht erforderlich, weil in beiden Aufnahmen die einzelnen Knochen segmentiert werden können und so eine direkte Berechnung von Transformationen möglich ist. Im multi-modalen Fall ist die Segmentierung i.d.R. lediglich in einer Modalität – nämlich im CT – vorhanden. Die Transformationen werden dann üblicherweise durch bildbasierte rigide Registrierungen ermittelt.

**Unabhängige Registrierung:** In den meisten Fällen erfolgt die Registrierung der einzelnen Strukturen unabhängig voneinander. Al-Mayah et al. [119] registrieren Wirbel und Unterkiefer rigide (CT/CT), wobei wichtige Details, wie z.B. das Ähnlichkeitsmaß nicht genannt werden. Zur CT/MR-Registrierung der einzelnen Wirbel maximieren Huesmann et al. [120] Mutual Information Wirbel für Wirbel, wobei das Ähnlichkeitsmaß nur innerhalb des segmentierten Bereichs berechnet wird. Hu et al. [121] und Čech et al. [122] gehen sehr ähnlich vor. Der einzige Unterschied besteht darin, dass der Bereich, in dem die Mutual Information berechnet wird, um einige Millimeter vergrößert wird, so dass die kontrastreichen Regionen am Rande der Wirbel (Übergang zu anderem Gewebe) in den Registrierungsprozess einbezogen werden. Die Robustheit kann dadurch erhöht werden. Ist der Bereich klein genug gewählt, sind Ungenauigkeiten nicht zu erwarten, weil angenommen werden kann, dass in diesem Bereich die Transformation zum Knochen sehr ähnlich sein wird.

Der große Nachteil der unabhängigen Registrierung besteht darin, dass der Konvergenzbereich relativ gering ist. Bei schwacher Deformation ist dies sicherlich kein Problem, bei großen Deformationen hingegen sind einige Wirbel in beiden Bildern sehr weit voneinander entfernt, wodurch die Anforderung an die Initialisierung erheblich steigt.

**Registrierung unter Einbeziehung der relativen Lage:** Um diesem entscheidenden Nachteil etwas entgegen zu setzen, beziehen Papademetris et al. [126] zur Registrierung von Mäuseextremitäten, Du Bois d'Aische et al. [117] zur Registrierung des Kopf-Hals-Bereichs, Martín-Fernández [128] zur Registrierung von Röntgenbildern von Händen und Baiker et al. [125] zur Registrierung von Mäuseskeletten die relative Lage der starren Strukturen zueinander in den Registrierungsprozess ein.

In [126] wird jeder Übergang zwischen zwei benachbarten Knochen durch ein Gelenk mit zwei Freiheitsgraden modelliert. Die einzelnen Transformationen werden nacheinander geschaltet. Die Transformation eines Punktes  $x$  des linken Knöchels im Referenzbild in den entsprechenden Punkt  $x'$  im



Modellbild ist bspw. durch

gegeben. Die Konkatenation von Transformationen dient als globales Vorwissen und verringert zum einen die Anzahl an Freiheitsgraden und vergrößert zugleich den Konvergenzbereich. Die  $N = 6 + 2n$  Parameter (global rigide,  $n$  Gelenke mit je 2 Parametern) werden gemeinsam mit Hilfe des *Robust-Point-Matching-Algorithmus* [135] bestimmt, welcher abwechselnd Punktkorrespondenzen zwischen extrahierten Knochenoberflächenpunkten bestimmt und die Transformationsparameter abschätzt.

Da in [128] die Landmarken aus beiden Aufnahmen der Hand extrahiert werden, können die Transformationsmatrizen der einzelnen Handknochen direkt aufgrund von den korrespondierenden Punkten berechnet werden. Ähnliches geschieht auch in [125]. Der zuvor beschriebene hierarchische artikulierte Atlas kann auf mehrere Bilder angewendet werden. Sind die Transformationen der einzelnen Elemente zum Atlas durch die Anpassung erst einmal bekannt, können auch die relativen Transformationen der Knochen zwischen den Aufnahmen direkt berechnet werden.

Ausgehend von den rigiden Transformationen der einzelnen rigiden Strukturen  $T_1, \dots, T_n$  soll nun ein dichtes Deformationsfeld, d.h. jedem Voxel des Referenzbildes wird genau ein Deformationsvektor zugeordnet, bestimmt werden. Dazu gibt es drei Möglichkeiten: i) Die Transformationen dienen als Initialisierung für eine deformierbare Registrierung und ii) die Transformationen sind die treibenden Kräfte einer Finite-Elemente-Methode und iii) das Deformationsfeld wird durch Interpolation der einzelnen rigiden Transformationen bestimmt.

---

56

Zur Registrierung von Mäuseextremitäten verwenden Suh et al. [127] die multi-rigide Registrierung der Knochen gemeinsam mit einer Registrierung der Hautoberfläche als Initialisierung für den Demons-Algorithmus [58], welcher das Weichteilgewebe registriert. Li et al. [124] verwenden zur Registrierung des Weichteilgewebes von Ganzkörper-CT-Aufnahmen den in [137] vorgestellten auf adaptive radiale Basisfunktionen basierenden ABA-Algorithmus. Dieser maximiert die normalisierte Mutual-Information und sorgt durch die Verwendung lokaler Regularisierung für plausible Transformationen.

Im Wesentlichen kann diese Art der Registrierung nicht mehr als multi-rigide, sondern mehr als frei deformierbar aufgefasst werden, welches die üblichen Nachteile mit sich bringt: hohe Komplexität und die Anfälligkeit für eine Anpassung an falsche Strukturen, wie z.B. Bildartefakte. Darüber hinaus werden nicht nur eine Änderung der Position, sondern auch etwaige Änderungen in der Anatomie ausgeglichen. Im Gegensatz zum Inter-Subjekt-Fall ist dies allerdings im Intra-Subjekt-Fall nicht immer gewünscht, z.B. wenn es darum geht, durch Registrierung positionskorrigierte Bilder einer Zeitreihe im Hinblick auf Anatomieänderungen zu vergleichen.

**Finite-Elemente-Methoden:** Zur Registrierung des Kopf-Hals-Bereichs verwenden Al-Mayah et al. [119] und Du Bois d'Aische et al. [117] Finite-Elemente-Methoden, welche durch die bekannten Transformationen der rigiden Strukturen getrieben werden. Bei Ersterem kommt dazu die von Brock et al. vorgestellte Methode MORFEUS [138] zum Einsatz. Bei Letzterem wird ein linear elastisches Modell angenommen. Zunächst wird ein Mesh aus Tetraeder generiert, welches die Größe der einzelnen Elemente nach der Wichtigkeit der Region bestimmt (z.B. hohe Auflösung um die rigiden Bereiche). Die Berechnung der Deformation  $\mathbf{u}$  für jeden einzelnen Knoten des Meshes wird durch ein dünnbesetztes lineares Gleichungssystem der Form

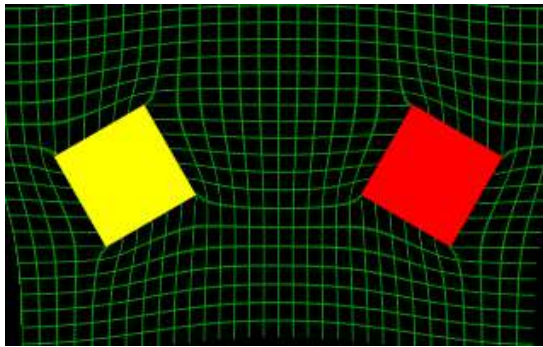
$$K\mathbf{u} = -\mathbf{f} \quad (3.14)$$

gelöst, wobei  $\mathbf{f}$  den durch die rigiden Transformationen induzierten externen Kräften und  $K$  einer Steifigkeitsmatrix entspricht. Die Deformation für jeden einzelnen Voxel wird dann durch Interpolation der Deformationen der umliegenden Meshknoten bestimmt.

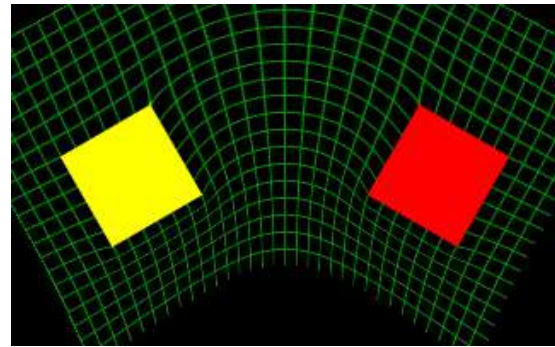
Wird das Gewebe exakt modelliert, kann sicherlich eine sehr hohe Registrierungsgenauigkeit im Weichteilgewebe erzielt werden. Dies erfordert allerdings eine umfangreiche Segmentierung von allen relevanten Strukturen mit unterschiedlichen Gewebeeigenschaften. Diese ist in der Regel nicht gegeben. Wird das gesamte Gewebe als homogen angenommen, entspricht dies im Wesentlichen einer Interpolation, wobei allerdings die Laufzeit durch die Meshgenerierung und das Lösen des Gleichungssystems eine hohe Rechenzeit mit sich bringt.

**Interpolationsschemata:** Durch zufällige Abtastung der Ränder der einzelnen rigiden Strukturen bestimmen Huesman et al. [120] zur Registrierung der Wirbelsäule korrespondierende Landmarken. Diese dienen dann als Stützpunkte für Thin-Plate-Splines. Verschiebungen der einzelnen rigiden Transformationen verbreiten sich so in das umliegende Weichteilgewebe.

Little et al. argumentierten in [54] allerdings, dass bei der Verwendung von Landmarken Deformationen in den eigentlich als starr betrachteten rigiden Strukturen auftreten können. Um dieses Problem zu umgehen, schlugen sie ein Interpolationsschema vor, welches mit Hilfe der Distanztransformation der rigiden Strukturen Gewichte  $w_i(\mathbf{x})$  berechnet, mit denen Linearkombination  $T(\mathbf{x}) = \sum_{\forall i} w_i T_i$



(a) Keine Fortpflanzung von Rotationen bei der Interpolation des durch eine Rotation entstandenen Deformationsfeldes



(b) Die Interpolation aller Anteile der Transformationen ermöglicht, dass auch Rotationen in die umliegenden Strukturen propagiert werden.

Abbildung 3.16.: Interpolation zweier um  $30^\circ$  gegensätzlich rotierter starrer Körper durch unterschiedliche Interpolationsstrategien.

der einzelnen rigiden Transformation  $T_i$  berechnet werden. So werden also nicht nur Translations-, sondern auch die Rotationskomponenten der rigiden Transformationen in das Weichteilgewebe propagiert (vgl. Abbildung 3.16). Dieses Schema hat sich weit verbreitet und wird unter anderem auch zur Registrierung der Wirbelsäule angewendet [122].

#### 3.2.5. Zusammenfassung

Zur Registrierung von Lymphknoten im Kopf-Hals-Bereich können lokal rigide Verfahren nicht ausschließlich eingesetzt werden, da aufgrund unterschiedlicher Positionen der Wirbelsäule der Konvergenzbereich zu gering ist. Daher muss der gesamte Kopf-Hals-Bereich registriert werden. Hierbei kommt es besonders auf die Wahl der Transformationsklasse an. Multi-rigide Verfahren bieten einen guten Kompromiss zwischen der Flexibilität auf der einen und Robustheit auf der anderen Seite. In dem betrachteten Intra-Subjekt-Fall sind innerhalb der Knochen keine Deformationen zu erwarten. Daher können die einzelnen Knochen als für sich rigide betrachtet werden, weswegen multi-rigide Verfahren in diesem Fall besonders geeignet sind. Die Registrierung gliedert sich in 3 Schritte: Segmentierung der Knochen, Registrierung der Knochen und Bestimmung des Deformationsfeldes im Weichteilgewebe.

Da insbesondere benachbarte Wirbel eine sehr ähnliche Form und ein sehr ähnliches Aussehen haben, muss bei der Segmentierung der einzelnen Knochen ein lageorientiertes Verfahren zur Anwendung kommen. Durch die große Beweglichkeit der Wirbelsäule können einzelne Bildbereiche in den unterschiedlichen Aufnahmen sehr weit voneinander entfernt sein. Um solche Aufnahmen dennoch erfolgreich registrieren zu können, ist eine individuelle Registrierung der einzelnen Knochen nicht ausreichend. Die Einbeziehung von Nachbarschaftsrelationen kann Abhilfe schaffen. Größte Robustheit liefern dabei Verfahren mit globaler Regularisierung.

Bei der Bestimmung der Deformation des Weichteilgewebes haben deformierbare Registrierungsmethoden aufgrund der hohen Parameteranzahl eine hohe Komplexität und sind sehr anfällig für Bildstörungen. FEMs haben üblicherweise eine höhere Komplexität als Interpolationsmethoden und

| Verfahren          | Jahr | Intra-Subjekt | Voll automatisches Verfahren | Multi-Modal möglich | Multi-Rigide | Automatische Seg. benötigter Teile | Abhängige Reg. der rigiden Teile | Behandlung des Weichteilgewebes | Anwendung auf Kopf-Hals | Quant. Evaluation der Weichteile |
|--------------------|------|---------------|------------------------------|---------------------|--------------|------------------------------------|----------------------------------|---------------------------------|-------------------------|----------------------------------|
| Hill[97]           | 1994 | ✓             | ✗                            | ✓                   | ✗            | -                                  | -                                | -                               | ✓                       | ✓                                |
| Little[54]         | 1996 | ✓             | ✗                            | -                   | ✓            | ✗                                  | ✗                                | INT                             | ✓                       | ✗                                |
| Studholme[99]      | 1996 | ✓             | ✓                            | ✓                   | ✗            | -                                  | -                                | -                               | ✓                       | ✓                                |
| Huesmann[120]      | 2003 | ✓             | ✗                            | ✓                   | ✓            | ✗                                  | ✗                                | INT                             | ✗                       | ✗                                |
| Hu[121]            | 2004 | ✓             | ✗                            | ✓                   | ✓            | ✗                                  | ✗                                | -                               | ✗                       | ✗                                |
| Papademetris[126]  | 2005 | ✓             | ✗                            | ✗                   | ✓            | ✗                                  | ✓                                | INT                             | ✗                       | ✗                                |
| Cech[122]          | 2006 | ✓             | ✓                            | ✓                   | ✓            | ✓                                  | ✗                                | INT                             | ✗                       | ✗                                |
| Li[124]            | 2006 | ✓             | ✓                            | ✗                   | ✓            | -                                  | ✓                                | REG                             | ✗                       | ✗                                |
| Urschler[118]      | 2006 | ✓             | ✓                            | ✗                   | ✓            | ✓                                  | ✗                                | INT                             | ✓                       | ✗                                |
| Baiker[125]        | 2007 | ✓             | ✓                            | ✗                   | ✓            | ✓                                  | ✓                                | -                               | ✗                       | ✗                                |
| D.B. d'Aische[117] | 2007 | ✓             | ✗                            | ✓                   | ✓            | ✗                                  | ✓                                | FEM                             | ✓                       | ✓                                |
| Commowick[68]      | 2008 | ✗             | ✗                            | ✗                   | ✗            | -                                  | -                                | REG                             | ✓                       | ✓                                |
| Han[112]           | 2008 | ✗             | ✓                            | ✗                   | ✗            | -                                  | -                                | REG                             | ✓                       | ✓                                |
| Klinder[133]       | 2008 | -             | ✓                            | -                   | ✓            | ✓                                  | ✓                                | -                               | ✗                       | -                                |
| Li[123]            | 2008 | ✓             | ✗                            | ✗                   | ✓            | ✗                                  | ✗                                | REG                             | ✗                       | ✗                                |
| Papastavrou [111]  | 2009 | ✓             | ✓                            | ✗                   | ✗            | -                                  | -                                | -                               | ✓                       | ✗                                |
| Gorthi[70]         | 2009 | ✗             | ✗                            | ✗                   | ✗            | -                                  | -                                | REG                             | ✓                       | ✓                                |
| M.-Fernández [128] | 2009 | ✓             | ✓                            | ✗                   | ✓            | ✓                                  | ✓                                | INT                             | ✗                       | ✓                                |
| Al-Mayah[119]      | 2010 | ✓             | ✗                            | ✗                   | ✓            | ✗                                  | ✗                                | FEM                             | ✓                       | ✓                                |
| Gering[109]        | 2010 | ✗             | ✓                            | ✗                   | ✗            | -                                  | -                                | REG                             | ✓                       | ✓                                |
| Hollensen[106]     | 2010 | ✗             | ✓                            | ✗                   | ✗            | -                                  | -                                | REG                             | ✓                       | ✓                                |
| Ramus[107]         | 2010 | ✗             | ✓                            | ✗                   | ✗            | -                                  | -                                | REG                             | ✓                       | ✓                                |
| Teng[71]           | 2010 | ✗             | ✓                            | ✗                   | ✗            | -                                  | -                                | REG                             | ✓                       | ✓                                |
| Yang[110]          | 2010 | ✗             | ✓                            | ✗                   | ✗            | -                                  | -                                | REG                             | ✓                       | ✓                                |
| Suh[127]           | 2011 | ✓             | ✓                            | ✗                   | ✓            | -                                  | ✓                                | REG                             | ✗                       | ✓                                |

Tabelle 3.2.: Zusammenfassung von Eigenschaften der in diesem Kapitel betrachteten Registrierungsmethoden. Nicht relevante/beschriebene Merkmale sind durch - gekennzeichnet. Die Behandlung des Weichteilgewebes kann entweder durch bildbasierte Registrierung (REG), Finite-Elemente-Methoden (FEM) oder Interpolation (INT) bestimmt werden.

benötigen evtl. physikalische Materialeigenschaften. Etwaige Ungenauigkeiten von Interpolationsmethoden können durch anschließende lokale Registrierung der Lymphknoten kompensiert werden.

Da, wie in Tabelle 3.2 zu sehen ist, kein existierendes Verfahren allen hier motivierten Kriterien genügt, besteht Bedarf an der Entwicklung eines neuartigen Verfahrens. Die größte Schwachstelle von in der aktuellen Literatur beschriebenen multi-rigiden Methoden liegt in der robusten Segmentierung der einzelnen rigiden Strukturen, welche bisher meist Interaktion durch den Anwender erfordern. Darüber hinaus kann der Bereich aller möglichen Transformationen durch die rigorose Einbringung von globaler Regularisierung weiter eingeschränkt werden. Bei existierenden Verfahren geschieht dies meist gar nicht und falls doch – wie z.B. in [125] und [117] – kann dies zur Erhöhung der Robustheit noch weiter ausgebaut werden.



## 4. Methoden

Aufgrund der in Kapitel 1 motivierten klinischen Bedeutung und der in Kapitel 3 identifizierten Lücken im heutigen Stand der Forschung beschreibt dieses Kapitel nun die im Rahmen der Arbeit entwickelten neuartigen Methoden zur Segmentierung und Registrierung von Lymphknoten. Die Evaluierung der Methoden erfolgt anschließend separat in Kapitel 5.

### 4.1. Segmentierung von Lymphknoten in CT-Datensätzen

Mit Blick auf die in Abschnitt 2.1 und 3.1 erarbeiteten Anforderungen, ist eine erfolgreiche Segmentierung von Lymphknoten keine triviale Aufgabe. Die größten Herausforderungen sind die hohe Varianz an Größe, Form und umliegendem Gewebe sowie das eventuelle Vorhandensein von Pathologien und der geringe Kontrast zu benachbarten Strukturen in CT-Aufnahmen. Keines der in der Literatur vorgestellten Verfahren betrachtet all dies gleichzeitig. Im Folgenden werden zunächst einige Erweiterungen des modellbasierten Ansatzes von Dornheim et al. [81] vorgestellt. Da auch diese nicht zu den gewünschten Ergebnissen führen, wird anschließend ein Radialstrahl-basierter Ansatz vorgestellt. Die Machbarkeit wird zunächst durch die Segmentierung von Lymphknoten aus 2D-CT-Schichten gezeigt. Darauf aufbauend wird dieser Ansatz dann zu einer echten 3D-Segmentierung erweitert. Besonderes Augenmerk wird der Auswahl der Kostenfunktion und der Bestimmung der dazugehörigen Parameter beigemessen, weil diese einen sehr großen Einfluss auf das Segmentierungsergebnis haben, in der Literatur dennoch meist ad hoc bestimmt werden. Als Initialisierung ist ein einzelner Saatpunkt in der Nähe des Objektschwerpunkts gegeben.

#### 4.1.1. Größeninvariantes Masse-Feder-Modell

Die erste im Rahmen dieser Arbeit entwickelte und in [8] veröffentlichte Methode zur Lymphknoten-segmentierung ist eine Erweiterung des Masse-Feder-Modells [47][81]. Prinzipiell erfüllt diese Methode viele der gewünschten Eigenschaften, insbesondere ist sie die einzige in der Literatur zu findende Methode, die auch Lymphknoten mit zentralem nekrotischem Gewebe explizit unterstützt. Durch die lokale Modellanpassung ergibt sich jedoch ein entscheidender Nachteil: Die starke Abhängigkeit von der Modellinitialisierung oder äquivalent von der unbekannten Größe des Lymphknotens. Wie bereits im vorangegangenen Kapitel beschrieben, wurde dieses Problem in [83] erkannt und durch einen Multi-Modell-Ansatz angegangen. Neben der Erhöhung der Komplexität, erhöht sich dadurch auch die Anzahl der freien Parameter und ein weiterer kritischer Schritt, die automatische Auswahl des am besten angepassten Modells, wird nötig. Anstatt eines Multi-Model-Ansatzes werden im Folgenden drei Erweiterungen, *relative Ruhelängen*, *ausgewogene Torsionskräfte* und eine *initiale Modellexpansion* eingeführt, die eine Größeninvarianz des Masse-Feder-Modells sicherstellen sollen.

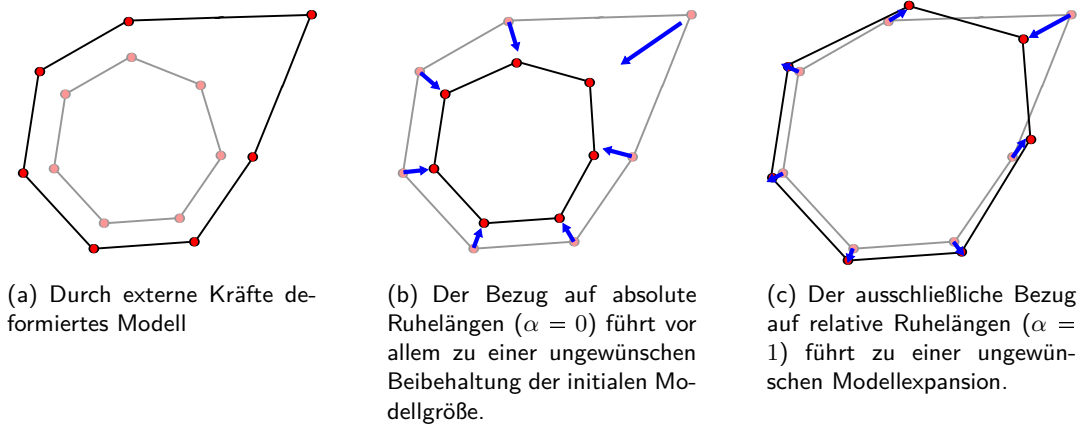


Abbildung 4.1.: Effekt der Federkräfte auf die Größe eines Masse-Feder-Modells, welches durch externe Kräfte deformiert wurde

**Relative Ruhelängen:** Die in [47] vorgeschlagenen Federkräfte  $\mathbf{F}_{i,j}$  zwischen den Massen  $i$  und  $j$  mit den Positionen  $\mathbf{s}_i$  und  $\mathbf{s}_j$  folgen der Newton'schen Mechanik und wirken so, dass die initiale Distanz zwischen den Massen  $l_{0ij} = |\mathbf{s}_j - \mathbf{s}_i|$  – die sog. Ruhelänge – möglichst erhalten bleibt:

$$\mathbf{F}_{i,j} = (|\mathbf{s}_j - \mathbf{s}_i| - l_{0ij}) \cdot \frac{\mathbf{s}_j - \mathbf{s}_i}{|\mathbf{s}_j - \mathbf{s}_i|}. \quad (4.1)$$

In dieser Gleichung entspricht der erste Faktor der Dehnung und der zweite Faktor der Kraftrichtung. Da  $l_{0ij}$  eine absolute Größe ist, ist dadurch auch eine globale Größe des Modells vorgegeben, von welcher durch das Wirken der externen Kräfte nur bedingt abgewichen werden kann. Eine Verringerung des Einflusses der Federkräfte kann diesen ungewünschten Effekt reduzieren, führt jedoch gleichzeitig zu einer Instabilität des Modells. Um dies zu verhindern wird Gleichung 4.1 dahingehend verändert, dass sich die Federkraft auf relative anstatt absolute Ruhelängen bezieht:

$$\mathbf{F}_{i,j} = (|\mathbf{s}_j - \mathbf{s}_i| - \lambda \cdot l_{0ij}) \cdot \frac{\mathbf{s}_j - \mathbf{s}_i}{|\mathbf{s}_j - \mathbf{s}_i|}, \quad (4.2)$$

wobei  $\lambda$  einem Skalierungsfaktor entspricht, welcher nach jeder Modellanpassungsiteration folgendermaßen neu bestimmt wird:

$$\lambda = \left( \frac{\sum_{\forall(i,j)} |\mathbf{s}_j - \mathbf{s}_i|}{\sum_{\forall(i,j)} l_{0ij}} \right)^\alpha. \quad (4.3)$$

Dabei sind  $(i, j)$  alle Paare von Massen, die miteinander durch eine Feder verbunden sind und  $\alpha$  ein Parameter, der eine ungewünschte Expansion verhindert. Stellt man sich bspw. eine einzelne Masse vor, welche durch eine externe Kraft stark in das Äußere gedrückt wird, während alle anderen Massen sich nur leicht nach Außen bewegen (siehe Abbildung 4.1(a)), sind zwei verschiedene extreme Arten der Kompensation durch Federkräfte denkbar: i) Bezug auf absolute Ruhelängen, die sämtliche Massen wieder in den Ursprungszustand zurück versetzen ( $\alpha = 0$ , siehe Abbildung 4.1(b)) und ii) Bezug auf relative Ruhelängen, wodurch die falsche Gesamtmodellgröße beibehalten wird ( $\alpha = 1$ , siehe Abbildung 4.1(c)). Ein  $\alpha \in ]0, 1[$  trifft einen Kompromiss zwischen den beiden Extremen, was zu einer stabileren Größe führt. Hier wurde  $\alpha = 0.95$  gewählt.

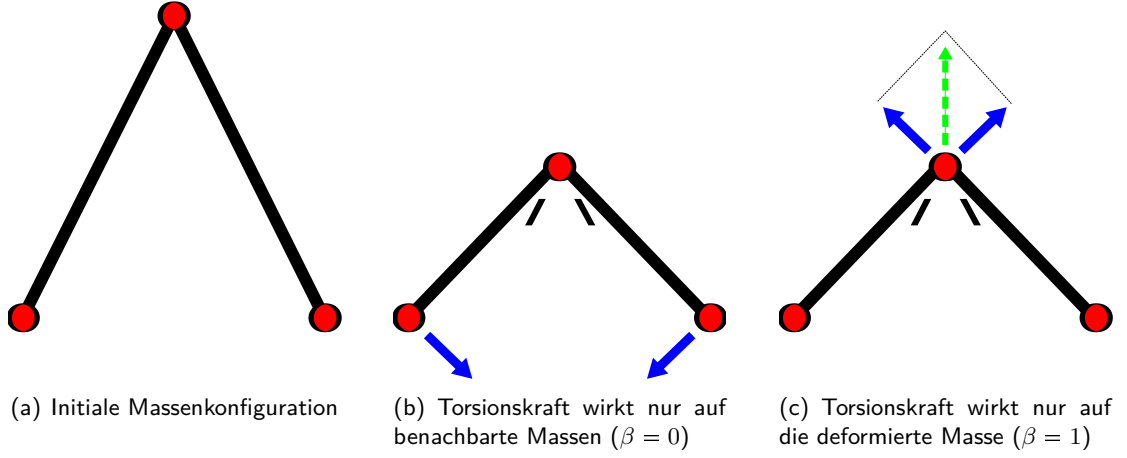


Abbildung 4.2.: Auswirkung der Anwendung von Torsionskräften auf unterschiedliche Massen

**Ausgewogene Torsionskräfte:** Zur Formerhaltung wurden in [47] Torsionskräfte eingeführt, die bei einer Positionsänderung einer Masse  $i$  auf die benachbarten Massen  $j$  wirken. Auf  $i$  selbst wirkt unmittelbar keine Torsionskraft, was den Formerhaltungsprozess in die Länge zieht. Daher werden in dieser Arbeit die Torsionskräfte dahingehend erweitert, dass sie sowohl auf die benachbarten Massen  $j$  als auch auf die Masse  $i$  selbst wirkt:

$$\mathbf{F}_{ij_j} = (1 - \beta) \cdot \mathbf{F}_{ij} \text{ und} \quad (4.4)$$

$$\mathbf{F}_{ij_i} = -\beta \cdot \mathbf{F}_{ij} . \quad (4.5)$$

$\beta \in [0, 1]$  ist dabei ein Parameter, der die in [47] beschriebene Torsionskraft  $\mathbf{F}_{ij}$  auf die Massen  $i$  und  $j$  aufteilt. Um eine Gleichverteilung zu erzielen, wird  $\beta = 0.5$  gewählt. In [47] wurde implizit  $\beta = 0$  gewählt. Abbildung 4.2 illustriert den erzielten Effekt.

**Initiale Modellexpansion:** Um zu verhindern, dass die Modellanpassung durch den Einfluss umliegender Strukturen fehlschlägt, muss das Modell komplett innerhalb des Lymphknotens initialisiert werden. Falls das Modell deutlich kleiner als der Lymphknoten ist, ist es wahrscheinlich, dass anfangs überhaupt keine externen Kräfte auf das Modell wirken, weil die Grauwerte an den Intensitätssensoren sich im gewünschten Bereich befinden und auch keine Bildgradienten im Konvergenzbereich der Gradientensensoren sind. Folglich kann sich das Modell nicht an die Lymphknoten anpassen. Aus diesem Grund wird eine initiale Modellexpansion eingeführt. Dies geschieht, indem jeder Masse  $i$  eine initiale Bewegung  $\mathbf{v}_{0_i}$  zugeordnet wird, welche eine konstante Geschwindigkeit  $v_0$  in Modellnormalenrichtung  $\mathbf{n}_i$  beschreibt:

$$\mathbf{v}_{0_i} = v_0 \cdot \mathbf{n}_i . \quad (4.6)$$

Dies führt zu einer balonartigen Expansion des Modells, welche durch den Dämpfungsfaktor  $d$  in der Anpassungsgleichung 3.2 im Laufe des Anpassungsprozesses verringert wird. Weiterhin verringert sich der Einfluss dieser Expansion sobald externe Kräfte des Lymphknotenrandes wirken.

#### 4.1.2. Radialstrahl-basierte 2D-Segmentierung von Lymphknoten

Wie in dem Evaluationskapitel 5.1.2 zu sehen sein wird, sind die gerade beschriebenen Erweiterungen des Masse-Feder-Modells zwar in der Lage den Einfluss der initialen Modellgröße auf das Segmentierungsergebnis zu verringern, das fundamentale Problem der lokalen Modellanpassung wird allerdings nicht ausgeräumt. Daher wird in den folgenden Abschnitten eine fundamental andere Methode entwickelt, welche auf Radialstrahlen basiert und eine globale Optimierung ermöglicht. Bevor allerdings ein 3D-Ansatz vorgestellt wird, soll in diesem Abschnitt zunächst die Machbarkeit anhand der Segmentierung von Lymphknoten in einzelnen 2D-CT-Schichten nachgewiesen werden. Dieses Verfahren wurde in [9] veröffentlicht. Die 2D-Segmentierung alleine ermöglicht bereits folgende Anwendungen:

- Die Segmentierung der Lymphknoten in einer einzelnen axialen CT-Schicht ermöglicht eine automatische Bestimmung der axialen Lymphknotendurchmesser, welche z.B. von den RE-CIST [26]-Kriterien benötigt werden. Neben einer Verringerung der Begutachtungsdauer wird auch die Inter- und Intra-Beobachter-Variabilität verringert, weil die Bestimmung der Lymphknotenbegrenzung immer durch genau denselben Algorithmus erfolgt.
- Eine 2D-Segmentierung eignet sich sehr gut zur Initialisierung eines modellbasierten 3D-Ansatzes mit lokaler Anpassungsstrategie. So könnte bspw. die in [79] manuelle grobe Konturierung in einer Schicht entfallen und durch ein automatisches robustes 2D-Segmentierungsverfahren ersetzt werden.

Auf die grundsätzliche Vorgehensweise von Radialstrahl-basierten Segmentierungsmethoden wurde bereits in Abschnitt 3.1.5 kurz eingegangen. In dem hier vorliegenden 2D-Fall werden  $n$  Strahlen um einen vorgegebenen Saatpunkt  $s$  gleichmäßig verteilt (siehe Abbildung 4.3(a)). Der  $j$ -te Punkt  $\mathbf{p}_{i,j}$  auf dem  $i$ -Strahl ergibt sich zu

$$\mathbf{p}_{i,j} = \mathbf{s} + j \frac{r_{\max}}{m} \begin{pmatrix} \cos\left(\frac{2\pi i}{n}\right) \\ \sin\left(\frac{2\pi i}{n}\right) \end{pmatrix}. \quad (4.7)$$

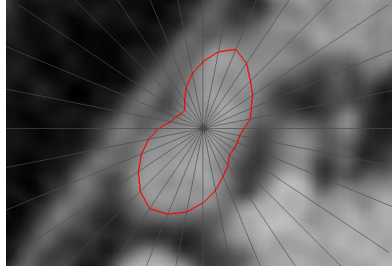
Ein  $n$ -dimensionaler Vektor von Punktindizes  $\mathbf{j}$  beschreibt einen geschlossenen Pfad um  $s$  und damit auch eine Segmentierung.  $r_{\max}$  beschreibt den maximal möglichen Radius und  $m$  gibt die Anzahl der Punkte eines Strahls. Das Energiefunktional aus Gleichung 3.4 reduziert sich in dem 2D-Fall zu

$$E(\mathbf{j}) = \sum_{i=0}^{n-1} (E_{\text{extern}}(\mathbf{p}_{i,j_i}) + E_{\text{intern}}(j_i, j_{i-1})) , \quad (4.8)$$

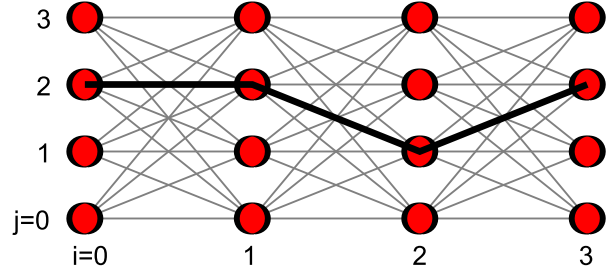
wobei die externe Energie  $E_{\text{extern}}(\mathbf{x})$  geringe Werte für Orte mit wahrscheinlicher Lymphknotengrenze annimmt und die interne Energie  $E_{\text{intern}}(j_i, j_{i-1})$  für lokale Formerhaltung sorgt. Um einen geschlossenen Pfad zu erhalten, wird  $j_{-1} := j_{n-1}$  definiert. Im Folgenden werden nun die einzelnen Energietermine motiviert und beschrieben. Alle angegebenen Parameter haben sich durch informelle Experimente als brauchbar erwiesen.

**Bildgradienten-Energieterm:** Am Rand eines Lymphknotens wird eine gewisse Bildgradientenstärke  $|\nabla I(\mathbf{x})|$  erwartet. Der Bildgradienten-Energieterm spiegelt dies wider und ergibt sich daher zu

$$E_{\text{Grad}}(\mathbf{x}) = w_{\text{Grad}} \cdot \max(0, g_{\max} - |\nabla(G_{\sigma} \star I(\mathbf{x}))|) , \quad (4.9)$$



(a) Diskrete geschlossene Kontur eines Lymphknotens (rot) entlang von Radialstrahlen (grau)



(b) Graph mit  $n = 4$  Strahlen und  $m = 4$  Radien mit allen möglichen Übergängen (grau) und einer möglichen Segmentierung (fett)

Abbildung 4.3.: Zugrundeliegende Datenstruktur der Radialstrahl-basierten 2D-Lymphknotensegmentierungsmethode

wobei  $g_{\max}$  einer maximalen Gradientenstärke,  $\nabla$  dem Gradientenoperator,  $\star$  dem Faltungsoperator und  $G_\sigma$  einem Gaußkernel mit dem Glättungsparameter  $\sigma = 1 \text{ mm}$  entspricht. Die Energie ist durch das Intervall  $[0, g_{\max}]$  begrenzt, um sehr starke Bildgradienten, wie sie z.B. im Übergang von Knochen zu Weichteilgewebe vorliegen, nicht zu bevorzugen. Die Gradientenkosten fließen ad hoc mit dem Gewichtungsfaktor  $w_{\text{grad}} = \frac{1}{12}$  in die Gesamtenergie ein.

**Intensitäts-Energieterm:** Mit Ausnahme von nekrotischem Gewebe, sollten sich alle Intensitätswerte im Inneren des Lymphknotens in einem festen Bereich  $[I_{\min}, I_{\max}]$  der Hounsfield-Skala befinden. Dieser Bereich wurde hier ad hoc auf  $[35, 160]$  festgelegt. Für einen Strahl ergibt sich der Intensitäts-Energieterm zu

$$E_{\text{Int}}(\mathbf{p}_{i,j}) = w_{\text{Int}} \sum_{k=l}^{j-1} \begin{cases} I(\mathbf{p}_{i,k}) - I_{\max} & ; I(\mathbf{p}_{i,k}) > I_{\max} \\ I_{\min} - I(\mathbf{p}_{i,k}) & ; I(\mathbf{p}_{i,k}) < I_{\min} \\ 0 & ; \text{sonst} \end{cases}, \quad (4.10)$$

wobei  $l = \min_{\forall k | I(\mathbf{p}_{i,k}) \in [I_{\min}, I_{\max}]} (k)$  der kleinste Radius ist, dessen Bildintensität sich in dem typischen Intervall befindet. Unter der Annahme, dass sich der Saatpunkt innerhalb von etwaigem zentralem nekrotischem Gewebe befindet, wird dadurch sichergestellt, dass dies keinen Einfluss auf die Energie hat. Ansonsten gilt, dass die Energie mit der Größe des Bereichs außerhalb des typischen Intensitätsbereichs und der absoluten Differenz der Intensitätswerte zu dem Bereich steigt. Dieser Energieterm ist maßgeblich an dem Ausschluss von Strukturen jenseits der Lymphknotengrenze verantwortlich. Abbildung 4.12 illustriert den Kostenverlauf beispielhaft entlang eines Strahls. Der Intensitäts-Energieterm fließt ad hoc mit dem Faktor  $w_{\text{Int}} = \frac{1}{6}$  in die Gesamtenergie ein.

**Formerhaltungs-Energieterm:** Um eine glatte Kontur zu erhalten, sollen ähnliche Radien von benachbarten Strahlen bevorzugt werden. Diese Art des lokalen Formwissens fließt in die interne Energie ein:

$$E_{\text{intern}}(k, j) = \left| \frac{r_{\max}}{m} (k - j) \right|^\alpha, \quad (4.11)$$

wobei  $\alpha = 8$  sehr unähnliche Radien überproportional stark bestraft.

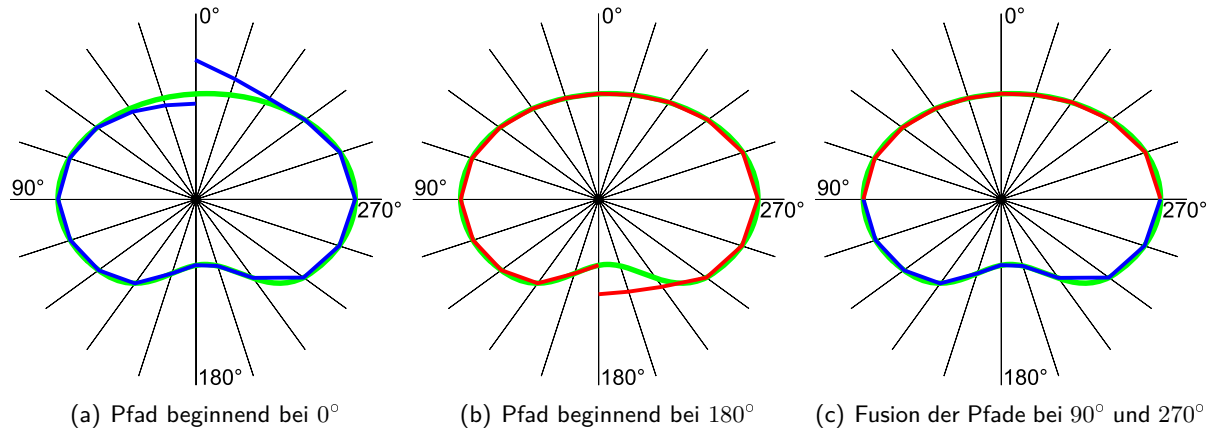


Abbildung 4.4.: Die Fusion zweier Pfade, die von gegenüberliegenden Seiten gestartet wurden, führt zu einem geschlossenen Pfad unter Beibehaltung der Komplexitätsklasse

**Optimierung:** Da das durch Gleichung 4.8 formulierte Energieminimierungsproblem eindimensional ist und nur direkt benachbarte Strahlen sich gegenseitig beeinflussen, können die einzelnen Radien als versteckte Zustände eines *Hidden Markov Models* angesehen werden. Die global optimale Sequenz kann effizient durch dynamische Programmierung gefunden werden [139]: Zunächst wird vorwärts über alle Strahlen  $i$  in dem in Abbildung 4.3(b) dargestellten Graphen iteriert. Dabei wird für jeden Radius  $j$  derjenige Radius  $k$  des Vorgängerstrahls  $i - 1$  mit der geringsten Summe an bisher kumulierter Energie plus der Formerhaltungsenergie des Übergangs  $E_{\text{intern}}(k, j)$  bestimmt und gemerkt. Zu dieser Energie wird die externe Energie  $E_{\text{extern}}(\mathbf{p}_{i,j})$  des aktuellen Knotens addiert, wodurch sich dessen kumulierte Energie ergibt. Wurde der letzte Strahl erreicht, wird der Radius mit der geringsten kumulierten Energie bestimmt, welcher der Ausgangspunkt zur Bestimmung des Pfades mit den geringsten Kosten ist. Die gesamte Sequenz von Radien wird nun durch Zurückverfolgen der im vorangegangenen Schritt gemerkten Übergängen bestimmt.

Die Gesamtkomplexität des Algorithmus ist durch  $\mathcal{O}(m^2 \cdot n)$  gegeben. Die Komplexität der Vorverarbeitung ist proportional zu der Gesamtanzahl an Pixeln in der Schicht.

Dieser Algorithmus funktioniert allerdings nur für lineare Pfade, die nicht geschlossen sind. Für geschlossene Pfade kann der optimale Pfad dennoch gefunden werden, indem der Algorithmus  $m$ -mal ausgeführt wird, wobei jedes Mal der Start- und End-Radius des Pfades auf einen anderen festen Wert gesetzt wird. Aus den  $m$  Pfaden wird dann durch einer erschöpfende Suche derjenige mit der geringsten Energie ausgewählt. Da dies aber die Komplexitätsklasse des Algorithmus von  $\mathcal{O}(m^2 \cdot n)$  auf  $\mathcal{O}(m^3 \cdot n)$  erhöht, wird der Algorithmus nur zweimal ausgeführt, einmal o.B.d.A. beginnend bei 0° und einmal auf der gegenüberliegenden Seite bei 180° (siehe Abbildung 4.4). In der Praxis hat sich gezeigt, dass die durch diese Heuristik gefundenen Pfade, mit Ausnahme von ein paar Strahlen in der Nähe von 0 bzw 180°, identisch zu dem optimalen geschlossenen Pfad sind. Eine Fusion der beiden Pfade bei 90° und 270° führt zu keinem Verlust an Segmentierungsgenauigkeit.

### 4.1.3. Radialstrahl-basierte 3D-Segmentierung von Lymphknoten

Wie die Evaluation des gerade beschriebenen Radialstrahl-basierten 2D-Lymphknotensegmentierungsansatzes in Abschnitt 5.1.3 zeigen wird, ist diese Art von Verfahren besonders durch den Einsatz eines globalen Optimierungsverfahrens vielversprechend und es lohnt sich daher das Verfahren auf den 3D-Fall zu erweitern. Anstatt durch einen geschlossenen Pfad ist eine Segmentierung nun durch eine geschlossene Oberfläche um einen Saatpunkt gegeben (vgl. Abschnitt 3.1.5 und siehe Abbildung 3.9). Die Radialstrahlen sind nicht in einer Ebene in gleichmäßig verteilten Winkeln angeordnet (siehe Abbildung 4.3(a)), sondern gleichmäßig auf einer Kugeloberfläche. Dies kann bspw. durch eine reguläre Tessellierung einer Kugel erzielt werden. Das Nichtvorhandensein einer natürlich gegebenen Reihenfolge, und die damit verbundene zyklische Nachbarschaft, führt allerdings zu einem entscheidenden Unterschied: Eine global optimale Lösung kann nicht mehr durch dynamische Programmierung gefunden werden, sondern komplexe Optimierungsverfahren sind nötig. Sie können eine global optimale Lösung jedoch nur annähernd finden.

Deutlich wichtiger als das Optimierungsverfahren ist in der Regel aber die gewissenhafte Auswahl einer für den speziellen Anwendungsfall geeigneten Kostenfunktion [140] und deren Parameter. Eine populäre generische Methode ist eine statistische Modellierung von auftretenden Grauwertprofilen an der Objektgrenze aus repräsentativen Daten, für die eine Referenzsegmentierung vorliegt. Die Kosten an der Stelle  $x$  ergeben sich dann durch ein Distanzmaß zwischen dem tatsächlichen Profil an der Stelle  $x$  und dem gelernten Profil. Eine solche Vorgehensweise setzt jedoch eine gewisse Konsistenz der Grauwerte in der Umgebung um die Objektgrenze voraus. Bei Lymphknoten mag dies für das Innere gegeben sein, für das Äußere können jedoch keine Annahmen getroffen werden (vgl. Abschnitt 2.1.1). Daher kann das Lernen von Grauwertprofilen nicht eingesetzt werden und die Kostenfunktion kann nur auf elementaren Annahmen beruhen:

- Diskontinuität an der Objektgrenze,
- Bekannte Beschaffenheit im Inneren des Lymphknotens.

Da die Diskontinuitäten an der Objektgrenze in vielen Fällen örtlich nicht sehr stark ausgeprägt sind und auch die Beschaffenheit im umliegenden Gewebe oftmals der von Lymphknoten gleicht, ist eine rein bildabhängige Kostenfunktion nicht ausreichend und es muss Formwissen eingebracht werden. Wie auch schon im 2D-Fall soll dies zunächst durch die Annahme geschehen, dass benachbarte Strahlen eine ähnliche Distanz zum in der Nähe des Schwerpunkts liegenden Saatpunkts aufweisen. Dies hat zur Folge, dass der prinzipbedingt auf sternkonvexe Formen mit dem Saatpunkt im Zentrum beschränkte Formenraum noch weiter auf Formen mit glatten Oberflächen eingeschränkt wird. Die in Abbildung 2.3 dargestellten Formen von Lymphknoten erfüllen diese Eigenschaften.

Im folgenden Abschnitt werden zunächst Optimierungsverfahren vorgestellt, welche eine Belegung für die Strahlen (ein bestimmter Radius für jeden Strahl) finden, so dass ein Kompromiss aus Bildinformation auf der einen und lokaler Forminformation auf der anderen Seite entsteht. Im Anschluss daran werden möglicherweise geeignete bildgradienten- und intensitätsbasierte Kostenfunktionen vorgestellt und deren Parameter durch ein datengetriebenes Verfahren bestimmt. Darauf aufbauend wird die geeignetste Kombination an Kostenfunktionen ausgewählt und weitere Parameter, wie z.B. die optimale Anzahl an Radialstrahlen, werden bestimmt. Abschließend werden mögliche Erweiterungen beschrieben. Grundlage für den gesamten Prozess dient die in Abschnitt 2.1 kurz erwähnte und in Abschnitt 5.1.1 ausführlich beschriebene repräsentative Menge von 49 Lymphknoten aus CT-Aufnahmen, für die eine Referenzsegmentierung gegeben ist. Um eine unabhängige Evaluierung der

Methode zu ermöglichen, kommt dreifache Kreuzvalidierung für sämtliche datenabhängige Parameterbestimmungen zur Anwendung. Die Aufteilung der 49 Lymphknoten in Trainings- und Testmengen wird ebenfalls in Abschnitt 5.1.1 beschrieben. Dieses Verfahren wurde in [10] veröffentlicht.

### 4.1.4. Optimierungsverfahren und Formerhaltung

Aufgabe des Optimierungsverfahrens ist es, für jeden Strahl  $i \in [1, n]$  einen Radius  $r_i$  zu finden, so dass

- (i) eine bildabhängige Energiefunktion  $E_{\text{Bild}}$  möglichst gering ist und
- (ii) die entstehende Oberfläche lokal glatt ist.

Ein Oberflächenpunkt auf dem  $i$ -ten Strahl  $\mathbf{x}_i$  kann durch den Saatpunkt  $\mathbf{s}$ , die Richtung des Strahls  $\mathbf{d}_i$  und dem Radius  $r_i$  beschrieben werden:

$$\mathbf{x}_i = \mathbf{s} + r_i \mathbf{d}_i . \quad (4.12)$$

Um ein diskretes Optimierungsproblem zu erhalten, wird jeder Strahl gleichmäßig in  $m$  Teile unterteilt. Es wird nun nicht mehr nach einem Vektor von Radien  $\mathbf{r} \in \mathbb{R}^n$  sondern einem Vektor von Indizes  $\mathbf{l} \in [1, \dots, m]^n$  gesucht. Gleichung 4.12 ergibt sich zu

$$\mathbf{x}_i = \mathbf{s} + l_i \frac{r_{\max}}{m} \mathbf{d}_i , \quad (4.13)$$

wobei  $r_{\max}$  der maximal mögliche Radius und  $\frac{r_{\max}}{m}$  der Abstand zweier benachbarter Punkte auf dem Strahl ist. Zusammen mit der Anzahl an Strahlen bestimmt Letzteres die bestenfalls erreichbare Segmentierungsgenauigkeit.

Die bildabhängige Energiefunktion  $E_{\text{Bild}} : [1, \dots, m]^n \mapsto \mathbb{R}_0^+$  ergibt sich aus der Summe von unären Kosten  $C : \mathbb{R}^3 \mapsto \mathbb{R}_0^+$ , die allen möglichen Oberflächenpunkten zugeordnet werden:

$$E_{\text{Bild}}(\mathbf{l}) = \sum_{\forall i} C \left( \mathbf{s} + l_i \frac{r_{\max}}{m} \mathbf{d}_i \right) \quad (4.14)$$

$$= \sum_{\forall i} C_i(l_i) , \quad (4.15)$$

wobei  $C_i(j)$  lediglich eine vereinfachte Notation der Funktion  $C \left( \mathbf{s} + j \frac{r_{\max}}{m} \mathbf{d}_i \right)$  darstellt. Die unäre Kostenfunktion  $C$  hängt nur von dem durch den Strahl und den Radius gegebenen Ort und dem Bild ab, notwendigerweise nicht jedoch von anderen Strahlen und deren Radien. Sie sei zunächst unbekannt. Auf diese wird in späteren Abschnitten ausführlich eingegangen.

Die Modellierung der Glattheit kann auf unterschiedliche Art und Weise erfolgen und ist abhängig von der Art der Optimierung. Da das Durchprobieren aller möglichen Belegungen  $\mathbf{l}$  nicht effizient ist, muss ein Optimierungsverfahren verwendet werden, welches ggf. auf Optimalität verzichtet. Im Folgenden werden verschiedene lokale und globale Optimierungsverfahren samt geeigneter Formerhaltung vorgestellt.

**Lokale Optimierung mit Gradientenabstiegsverfahren:** Eine populäre und zugleich einfache Möglichkeit, die Formerhaltung zu modellieren, besteht darin, zu fordern, dass benachbarte



Strahlen ähnliche Radien haben. Dies wurde schon in [67] und bei der in Abschnitt 4.1.2 vorgestellten 2D-Segmentierung verwendet. Dort wurde dies durch Gleichung 4.11 sichergestellt. Die Folge ist, dass das Objekt als lokal kugelförmig angenommen wird. Im 3D-Fall ist diese Art der Formerhaltung durch die paarweisen Kosten  $C_{i,j}(l_i, l_j)$  benachbarter Strahlen  $i, j$  gegeben:

$$C_{i,j}(l_i, l_j) = \lambda \left| \frac{r_{\max}}{m} (l_i - l_j) \right|^\alpha . \quad (4.16)$$

Der Exponent  $\alpha \geq 1$  erlaubt eine überproportionale *Bestrafung* hoher Radiendifferenzen benachbarter Strahlen und  $\lambda$  entspricht den Kosten, die bei einer Radiendifferenz von 1 mm zugeordnet werden. Die Formenergie ergibt sich aus der Summe der paarweisen Kosten aus Gleichung 4.16 über alle Strahlen:

$$E_{\text{Form}}(\mathbf{l}) = \sum_{\forall i} \sum_{\forall j \in \mathcal{N}(i)} C_{i,j}(l_i, l_j) , \quad (4.17)$$

wobei  $\mathcal{N}(i)$  der Menge der benachbarten Strahlen des Strahls  $i$  entspricht.

Das zu minimierende Gesamtenergiefunktional setzt sich dann analog zur Gleichung 3.4 aus Bildenergie und Formerhaltungsenergie zusammen:

$$E(\mathbf{l}) = E_{\text{Bild}}(\mathbf{l}) + E_{\text{Form}}(\mathbf{l}) \quad (4.18)$$

$$= \sum_{\forall i} C_i(l_i) + \sum_{\forall i} \sum_{\forall j \in \mathcal{N}(i)} C_{i,j}(l_i, l_j) . \quad (4.19)$$

Da dieser Energieterm nur unäre ( $C_i$ ) und paarweise ( $C_{i,j}$ ) Kostenterme enthält, kann das Problem durch ein *Markov Random Field* (MRF) beschrieben werden, für das viele effiziente Energieminimierungsalgorithmen existieren. Da der eigentliche Wert für die Energie bei der Minimierung nicht relevant ist, muss Gleichung 4.19 nicht ausgewertet werden. Daher müssen mögliche numerische Probleme bei der Addition vieler kleiner Zahlen für eine hohe Anzahl von Strahlen und damit Summanden nicht beachtet werden.

Im einfachsten Fall erfolgt die Energieminimierung durch ein Gradientenabstiegsverfahren. Dabei wird ausgehend von einer Startbelegung  $\mathbf{l}^0$  immer mit kleinen Schrittweiten  $\Delta$  in negativer Richtung des Gradienten des Energiefunktionals gegangen, bis Konvergenz erreicht ist. Die Iterationsvorschrift der  $k$ -ten Iteration ist durch

$$\mathbf{l}^k = \mathbf{l}^{k-1} - \nabla E(\mathbf{l}^{k-1}) \quad (4.20)$$

gegeben. Die partiellen Ableitungen des Energiefunktionals  $\frac{\partial E}{\partial l_i}$  setzen sich aus der Summe der partiellen Ableitungen des Bildenergieterms und der des Formerhaltungsenergieterms zusammen. Wo hingegen letztere einfach analytisch bestimmt werden können, ist dies bei dem Bildenergieterm nicht möglich. Daher wird auf eine Approximation mittels finiter Differenzen zurückgegriffen:

$$\frac{\partial E_{\text{Bild}}}{\partial l_i} \approx \frac{E_{\text{Bild}}((l_1, \dots, l_i + 1, \dots, l_n)^T) - E_{\text{Bild}}((l_1, \dots, l_i - 1, \dots, l_n)^T)}{2} \quad (4.21)$$

$$= \frac{C_i(l_i + 1) - C_i(l_i - 1)}{2} . \quad (4.22)$$

Die initiale Belegung  $\mathbf{l}^0$  wird dadurch bestimmt, dass für jeden Strahl unabhängig derjenige Radius gesucht wird, welcher zu minimalen Kosten führt:

$$l_i^0 = \underset{k \in [1, \dots, m]}{\operatorname{argmin}} C_i(k) . \quad (4.23)$$

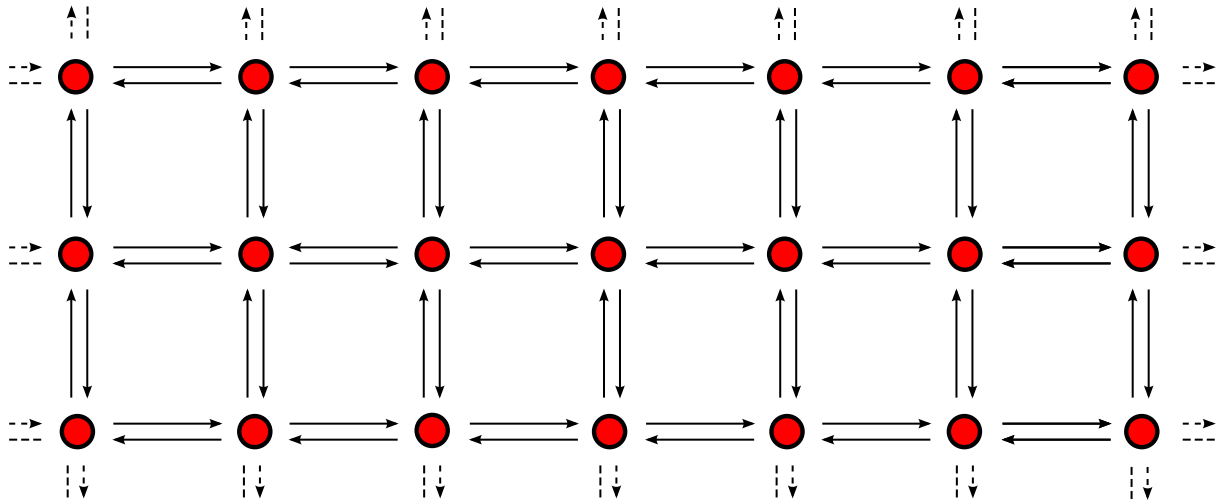


Abbildung 4.5.: Durch den Belief-Propagation-Algorithmus werden zwischen benachbarten Strahlen (rot) Nachrichten ausgetauscht. Die Belegung eines Strahls hängt sowohl von den bildbasierten Kosten des Strahls als auch von eingehenden Nachrichten ab.

Der große Vorteil dieses Optimierungsverfahrens ist das günstige Laufzeitverhalten. Da es sich allerdings um ein lokales Optimierungsverfahren handelt, hängt die erreichbare Segmentierungsgenauigkeit zum einen von der Initialisierung und zum anderen von dem Energiefunktional ab. Je näher die Initialisierung an dem globalen Minimum und je wenige lokale Minima die Zielfunktion aufweist, desto bessere Ergebnisse werden erwartet. Besonders durch das inhomogene umliegende Gewebe ist eine *gutartige* Zielfunktion allerdings nicht zu erwarten. Dennoch verwenden Barbu et al. in ihrem Radialstrahl-basierten Verfahren [67] ein Gradientenabstiegsverfahren zum Finden der Belegung.

**Globale Optimierung mit Belief-Propagation:** Ein Optimierungsverfahren, welches oft für Energien der Form aus Gleichung 4.19 verwendet wird und nicht die Nachteile eines lokalen Optimierungsverfahrens hat, ist *Belief-Propagation* [42] (BP).

Dabei wird jeder möglichen Belegung  $l_i \in [1, \dots, m]$  des Strahls  $i$  eine Wahrscheinlichkeit zugeordnet, die sich zunächst nur aus den unären Kosten  $C_i(l_i)$  ergibt. Abhängig von dieser Wahrscheinlichkeit sendet jeder Strahl zu seinen Nachbarstrahlen  $j \in \mathcal{N}(i)$  für jede mögliche Belegungskombination  $(l_i, l_j)$  eine Nachricht, die sich aus der eigenen Wahrscheinlichkeit und den durch die Formerhaltung gegebenen paarweisen Übergangskosten  $C_{i,j}(l_i, l_j)$  zusammensetzt. Aus den Nachrichten, die ein Strahl erhält und den aktuellen eigenen Wahrscheinlichkeiten werden letztere neu berechnet (vgl. Abbildung 4.5). Dieser Vorgang wird iterativ fortgesetzt. Abschließend wird für jeden Strahl  $i$  diejenige Belegung  $l_i$  mit der größten Wahrscheinlichkeit gewählt.

Sind die Nachbarschaften so gewählt, dass keine Zyklen entstehen, ist Optimalität garantiert. Ein Beispiel dafür ist das in Abschnitt 4.1.2 vorgestellte Radialstrahl-basierte 2D-Segmentierungsverfahren, wo BP äquivalent zu der dort verwendeten dynamischen Programmierung ist. In dem 3D-Fall treten allerdings Zyklen auf und Optimalität kann nicht garantiert werden. Dennoch erzielt BP auch dann in vielen Fällen sehr gute Ergebnisse.

BP bietet viele Vorteile. Eine Initialisierung ist nicht notwendig und die Wahrscheinlichkeit in einem

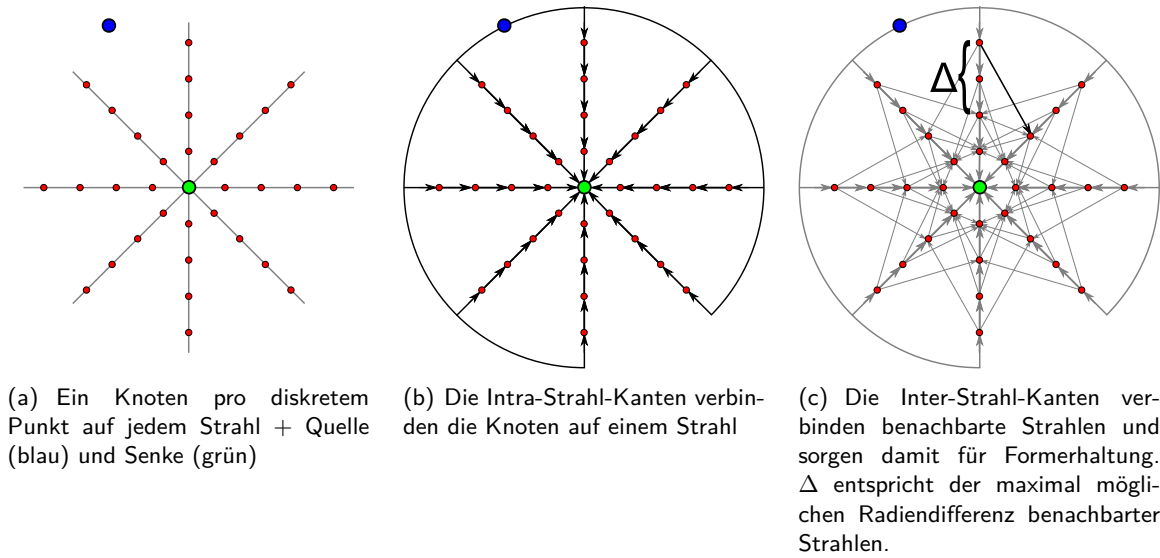


Abbildung 4.6.: Prinzipieller Ablauf der Grapherzeugung veranschaulicht für den 2D-Fall mit Formerhaltungsparameter  $\Delta = 2$

lokalen Minimum zu landen, welches weit entfernt von dem globalen Minimum ist, ist deutlich geringer als bei einem lokalen Optimierungsverfahren. Darüber hinaus eignet sich BP innerhalb eines Iterationsschritts sehr gut zur Parallelisierung. Durch den iterativen Charakter des Verfahrens können die zugrundeliegenden Kosten  $C_i(l_i)$  auch während der Berechnung geändert werden. Dies wird in der in Abschnitt 4.1.8 vorgestellten Erweiterung zur interaktiven Segmentierung ausgenutzt.

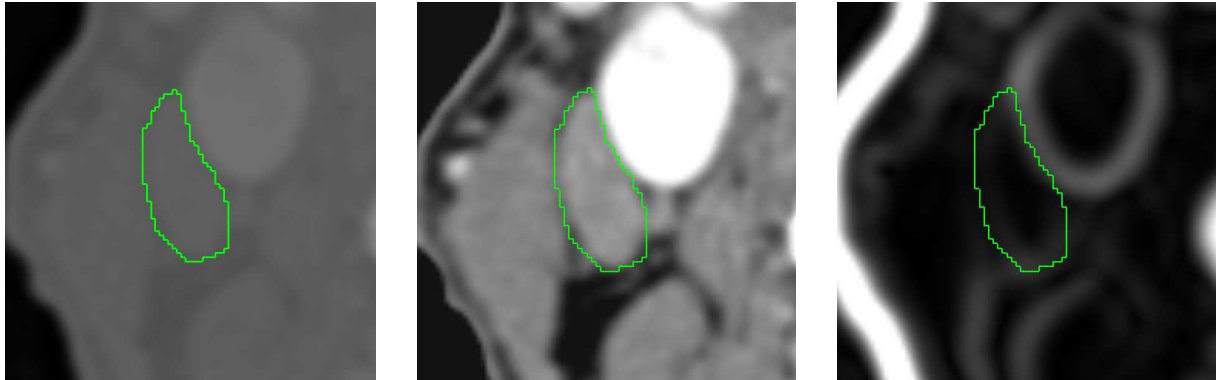
**Globale Optimierung mit Graph-Cut:** Ein weiteres globales Optimierungsverfahren, das sich im Kontext der Radialstrahl-basierten Segmentierung eignet, ist *Graph-Cut* [41] (GC). Es kann einen optimalen Schnitt und damit eine geschlossene Oberfläche zwischen dem Saatpunkt und den äußersten Punkten eines jeden Strahls finden. Diese Art der Oberflächenfindung wurde zunächst in [94] für allgemeine Oberflächen vorgestellt und anschließend in [91] und [93] zur Radialstrahl-basierten Segmentierung von Gliomen bzw. Lymphknoten verwendet (vgl. Abschnitt 3.1.5).

Der zugrundeliegende knotengewichtete gerichtete Graph besteht aus einem Knoten  $V_{i,j}$  für jeden Punkt  $j$  auf jedem Strahl  $i$  und zwei weiteren ausgezeichneten Knoten *Quelle* und *Senke* (siehe Abbildung 4.6(a)). Die Gewichte  $w$  der Knoten sind durch die Differenzen der Kosten der entsprechenden Punkte gegeben:

$$w_{i,j} = \begin{cases} C_i(j) & ; j = 1 \\ C_i(j) - C_i(j-1) & ; 2 \leq j \leq m \end{cases} \quad (4.24)$$

Die Kanten setzen sich aus *Intra-Strahl*- und *Inter-Strahl*-Kanten zusammen. Die Intra-Strahl-Kanten verbinden den  $j$ -ten Knoten auf jedem Strahl mit dem  $j-1$ -ten Knoten. Zusätzlich wird der Quellenknoten mit allen äußersten Knoten und der Knoten des Saatpunkts mit dem Senkenknoten verbunden (siehe Abbildung 4.6(b)).

Die Formerhaltung wird nicht mehr, wie bei den vorangegangenen Optimierungsverfahren, durch Gleichung 4.17 gegeben, sondern wird durch die Inter-Strahl-Kanten modelliert (siehe Abbildung 4.6(c)).



(a) In der Original-CT-Aufnahme wird nur ein kleiner Intensitätsbereich für Weichteilgewebe verwendet.

(b) Intensitätskorrigiert, alle Intensitäten außerhalb des Intervalls  $[-100, 260]$  HU werden herausgefiltert.

(c) Die Grenzen anderer Strukturen weisen eine deutlich höhere Bildgradientenstärke als die Lymphknotengrenzen auf

Abbildung 4.7.: Lymphknoten neben einem kontrastverstärkten Gefäß (rechts oben), Fettgewebe (darunter) und anderem Weichteilgewebe (links). Eine Referenzsegmentierung ist in grün eingezeichnet.

Der  $j$ -te Knoten eines jeden Strahls  $i$  wird dazu mit dem Knoten  $\max(1, j - \Delta)$  aller benachbarter Strahlen verbunden. Der Parameter  $\Delta$  bestimmt den Grad der Glattheit der entstehenden Oberflächen. Für  $\Delta = 0$  werden nur Kugeln um den Saatpunkt  $s$  erlaubt. Größere  $\Delta$  erhöhen die mögliche Flexibilität. Im Gegensatz zu der durch Gleichung 4.17 beschriebenen Formerhaltung werden nicht alle Übergänge benachbarter Strahlen erlaubt. Diejenigen, die erlaubt sind, haben allerdings alle dieselben Kosten.

Auf diesem Graph bestimmt dann der Min-Cut/Max-Flow Algorithmus von Boykov und Kolmogorov [141] denjenigen Schnitt zwischen Quelle und Senke, welcher zu den geringsten kumulierten Kosten unter Beachtung der geforderten Glattheit führt.

Auch dieses Verfahren benötigt keine Initialisierung. Die erhaltene Lösung ist optimal, allerdings wurde die weiche Formerhaltung aus Gleichung 4.17, welche eine geringe Radiendifferenz bevorzugt, durch eine harte Randbedingung ersetzt, welche maximal eine Radiendifferenz von  $\Delta$  zulässt.

### 4.1.5. Bildgradientenbasierte Kostenfunktionen

Nachdem in den vorangegangenen Abschnitten auf das generelle Segmentierungsverfahren, die Formerhaltung und mögliche Optimierungsverfahren eingegangen wurde, widmen sich die nächsten Abschnitte dem wohl wichtigsten, aber in der Literatur notorisch vernachlässigten Aspekt der bildabhängigen Kostenfunktion  $C : \mathbb{R}^3 \mapsto \mathbb{R}_0^+$ . Diese soll abhängig von dem zu segmentierenden Bild  $I : \mathbb{R}^3 \mapsto \mathbb{R}$  einem Ort  $\mathbf{x} \in \mathbb{R}^3$ , möglichst geringe Kosten zuordnen, falls dieser sich auf der Lymphknotenoberfläche befindet und andernfalls möglichst hohe Kosten. Wie bereits erläutert, kann die Kostenfunktion die Diskontinuität an der Objektgrenze und die bekannte Beschaffenheit im Inneren des Lymphknotens ausnutzen. Dieser Abschnitt beschäftigt sich mit Ersterem.

Diskontinuität an der Objektgrenze äußert sich durch eine lokal höhere Bildgradientenstärke als in

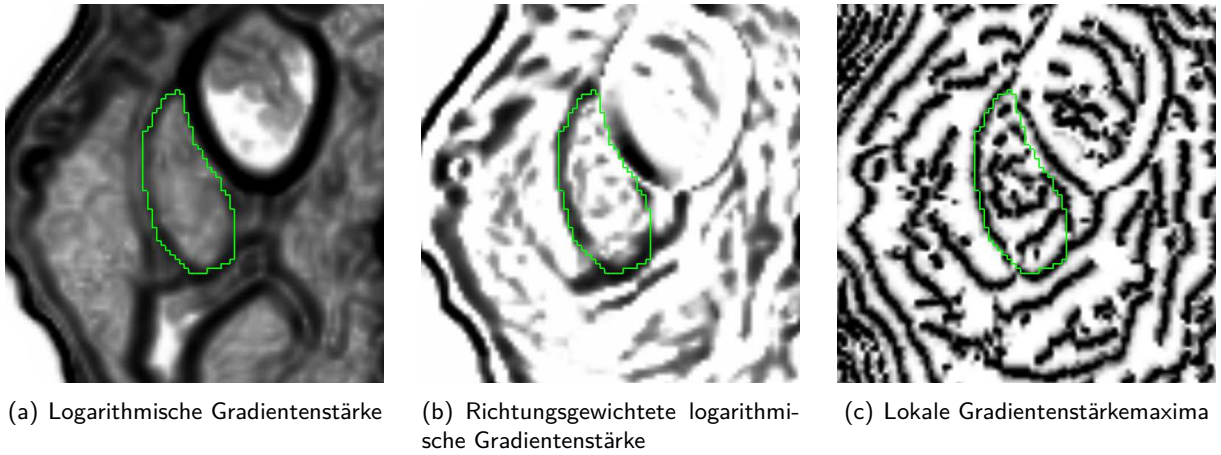


Abbildung 4.8.: Bildgradientenbasierte Kostenfunktionen im Vergleich

der Umgebung. Hohe Bildgradientenstärken treten natürlich auch an den Grenzen anderer Objekte auf und können dort – wie z.b. im Übergang von Knochen auf Weichteilgewebe – deutlich stärker ausfallen (siehe Abbildung 4.7(c)). Daher muss bei der Wahl der Kostenfunktion darauf geachtet werden, dass hohe Bildgradientenstärken zu geringen Kosten führen, sehr hohe Bildgradientenstärken allerdings nicht zu unverhältnismäßig niedrigen Kosten führen, um eine fälschliche Anpassung an andere, kontrastreichere umliegende Strukturen zu verhindern.

Da in CT-Aufnahmen der Grauwertbereich von Lymphknoten bekannt ist (siehe Abschnitt 2.1.3), kann dieses Wissen verwendet werden, um Bildgradienten in anderen Grauwertbereichen auszuschließen. Dies geschieht durch die Anwendung folgender Transferfunktion auf das Intensitätsbild  $I$ :

$$\bar{I}(\mathbf{x}) = \begin{cases} I_{\min} & ; I(\mathbf{x}) < I_{\min} \\ I_{\max} & ; I(\mathbf{x}) > I_{\max} \\ I(\mathbf{x}) & ; \text{sonst} \end{cases} . \quad (4.25)$$

Das Intervall  $[I_{\min}, I_{\max}]$  gibt den für den Übergang von Lymphknoten zu umliegenden Gewebe relevanten Bildintensitätsbereich vor. Dieses wird ad hoc auf das Intervall  $[-100, 260]$  festgelegt, so dass das Zentrum des Intervalls der durchschnittlichen Lymphknotenintensität entspricht (siehe Abbildung 2.5). Die Größe des Intervalls wurde so gewählt, dass alle innerhalb der Lymphknoten vorkommende Bildintensitäten und ein großer benachbarter Intensitätsbereich enthalten sind, nicht jedoch grundsätzlich verschiedene Gewebstypen wie Knochen oder Luft (siehe Abbildung 4.7(b)).

Besonders durch den Partialvolumeneffekt kann es sein, dass die durch die Referenzsegmentierung gegebene Kontur systematisch von lokalen Bildgradientenstärkenmaxima abweicht. Um dem entgegenzuwirken, wird ähnlich zu [82] der Parameter  $\delta$  eingeführt, welcher abhängig vom Saatpunkt  $\mathbf{s}$  die Position  $\mathbf{x}$  folgendermaßen korrigiert:

$$\bar{\mathbf{x}} = \mathbf{x} + \delta \cdot \frac{\mathbf{x} - \mathbf{s}}{|\mathbf{x} - \mathbf{s}|} . \quad (4.26)$$

Der Punkt wird also entlang eines Radialstrahls der Richtung  $\frac{\mathbf{x} - \mathbf{s}}{|\mathbf{x} - \mathbf{s}|}$  um den konstanten Wert  $\delta$  verschoben.

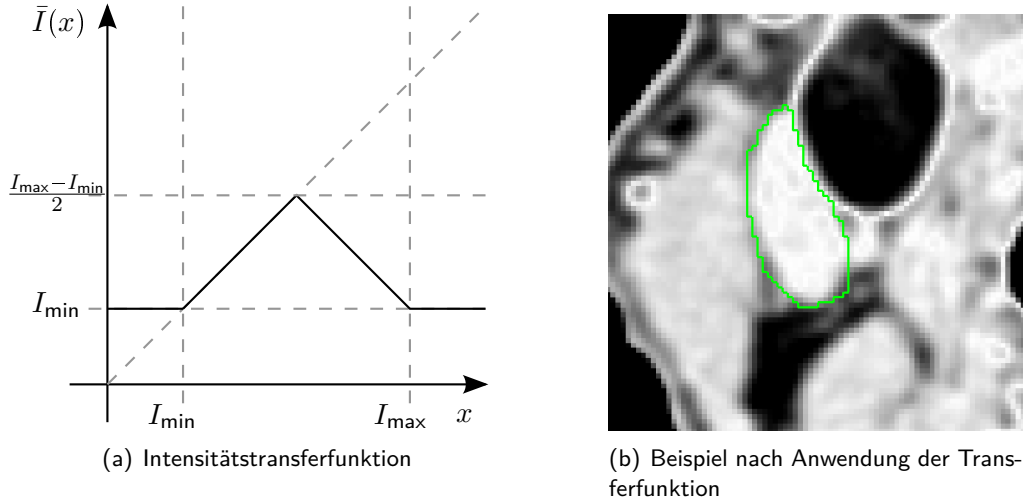


Abbildung 4.9.: Nach Anwendung der hier dargestellten Intensitätstransferfunktion weist Lymphknotengewebe die höchsten Bildintensitäten auf. Bildgradienten des transformierten Intensitätsbildes am Rand eines Lymphknotens zeigen folglich immer in Richtung des Lymphknotens.

**Logarithmische Gradientenstärke:** Eine einfache Möglichkeit, die gerade formulierten Anforderungen zu erfüllen, basiert auf der logarithmischen Gradientenstärke:

$$C_{\text{LogAbsGrad}}(\mathbf{x}) = C_{\text{max}} - \log(1 + |\nabla(G_\sigma * \bar{I}(\bar{\mathbf{x}}))|) . \quad (4.27)$$

Wie schon zuvor entspricht  $\nabla$  dem Gradientenoperator,  $\star$  dem Faltungsoperator und  $G_\sigma$  einem Gaußkern der Varianz  $\sigma^2$ . Die maximalen Kosten  $C_{\text{max}} = 5$  wurden so gewählt, dass  $C_{\text{LogAbsGrad}}(\mathbf{x}) > 0$  stets gilt. In homogenen Bildbereichen sind die Kosten maximal und mit zunehmender Bildgradientenstärke sinkt diese stetig, wobei durch die Logarithmisierung sehr starke Bildgradienten nur zu unwesentlich geringeren Kosten als etwas schwächere Bildgradientenstärken führen (siehe Abbildung 4.8(a)). Freie Parameter, die nicht ad hoc festgelegt werden können, sind der Bildglättungsparameter  $\sigma$  und die systematische Abweichung  $\delta$ .

**Richtungsgewichtete logarithmische Gradientenstärke:** Neben der Bildgradientenstärke ist es sinnvoll, auch die Bildgradientenrichtung einfließen zu lassen. Bei konturbasierten Verfahren geschieht dies üblicherweise durch Projektion der Bildgradientenrichtung entlang der Konturnormalen [142][88]. Bei Radialstrahl-basierten Methoden ist allerdings zum Zeitpunkt der Kostenberechnung keine Modellnormalenrichtung oder Ähnliches vorhanden. Ersatzweise kann aber auf die Richtung des Radialstrahls  $\mathbf{d} = \frac{\mathbf{x} - \mathbf{s}}{|\mathbf{x} - \mathbf{s}|}$  zurückgegriffen werden, welche jedoch die Konturnormalenrichtung nur annähert.

Ist das umliegende Gewebe dunkler (z.B. Fettgewebe), zeigt der Bildgradient an der Lymphknotengrenze in das Innere, bei hellerem umliegenden Gewebe (z.B. Knochen) in das Äußere. Um eine konsistente Gradientenrichtung unter Beibehaltung der Gradientenstärke zu erhalten, wurde in [82] die Plateau-Filterung (siehe Abbildung 3.6(a)) eingeführt, welche die Bildintensität so transformiert, dass Lymphknotengewebe die höchsten Bildintensitäten aufweist und somit alle Bildgradienten

an der Lymphknotengrenze immer in das Innere des Lymphknotens zeigen (siehe Abbildung 4.9). Um von diesem gewünschten Effekt bei der Gradientenkostenberechnung profitieren zu können, wird Gleichung 4.25 entsprechend modifiziert:

$$\bar{I}(\mathbf{x}) = \begin{cases} I_{\min} & ; I(\mathbf{x}) < I_{\min} \\ I_{\min} & ; I(\mathbf{x}) > I_{\max} \\ I_{\max} + I_{\min} - I(\mathbf{x}) & ; \frac{I_{\max} + I_{\min}}{2} < I(\mathbf{x}) \leq I_{\max} \\ I(\mathbf{x}) & ; \text{sonst} \end{cases} \quad (4.28)$$

Zur Berechnung der bildgradientenbasierten Kostenfunktion wird nun Gleichung 4.27 um eine Richtungsabhängigkeit erweitert:

$$C_{\text{RichtGrad}}(\mathbf{x}) = \begin{cases} C_{\max} & ; \angle(\mathbf{I}'(\mathbf{x}), -\mathbf{d}) \geq 90^\circ \\ C_{\max} - \log(1 + |\mathbf{I}'(\mathbf{x})|) \cdot \frac{\cos(\angle(\mathbf{I}'(\mathbf{x}), -\mathbf{d}))}{\cos \gamma} & ; \gamma < \angle(\mathbf{I}'(\mathbf{x}), -\mathbf{d}) < 90^\circ \\ C_{\max} - \log(1 + |\mathbf{I}'(\mathbf{x})|) & ; \angle(\mathbf{I}'(\mathbf{x}), -\mathbf{d}) \leq \gamma \end{cases} \quad (4.29)$$

wobei  $\mathbf{I}'(\mathbf{x}) = \nabla (G_\sigma \star \bar{I}(\bar{\mathbf{x}}))$  dem Bildgradienten entspricht. Eine ähnliche Fallunterscheidung wurde bereits in [88] verwendet und hat zur Folge, dass Bildgradienten, die bis zu einem Winkel von  $\gamma$  von der Richtung  $\mathbf{d}$  des Radialstrahls abweichen, mit vollem Gewicht in die Kostenberechnung einfließen. Dies ist insbesondere bei Radialstrahl-basierten Verfahren von Vorteil, da die Kontur in vielen Fällen nicht orthogonal zum Strahl verläuft. Bildgradienten mit einer höheren Winkelabweichung als  $\gamma$  werden entsprechend weniger stark gewichtet und Bildgradienten, die gar entgegengesetzt zum Strahl verlaufen, werden, wie homogene Strukturen auch, mit maximalen Kosten bewertet. Dadurch werden Übergänge des umliegenden Gewebes in andere Strukturen unterdrückt (siehe Abbildung 4.8(b)). Auch der Übergang von etwaigem zentralen nekrotischen Gewebe in das Lymphknotengewebe, wo eine beträchtliche Bildgradientenstärke vorliegen kann, wird unterdrückt, weil hier die Gradientenrichtung in entgegengesetzte Richtung zeigt.

Freie Parameter sind, wie bei der rein auf gradientenstärkebasierten Kostenfunktion auch, der Bildglättungsparameter  $\sigma$  und die systematische Abweichung  $\delta$ . Zusätzlich kommt der Winkel  $\gamma$  hinzu.

**Lokale Gradientenstärkemaxima:** Die bisher vorgestellten bildgradientenbasierten Kostenfunktionen hängen nach wie vor von der Bildgradientenstärke ab, obwohl es keine theoretische Rechtfertigung dafür gibt, dass ein stärkerer Bildgradient eher als ein schwächerer Bildgradient an der Objektgrenze eines Lymphknotens auftritt. Einen interessanten Ansatz, um komplette Unabhängigkeit von der Bildgradientenstärke zu erreichen, wurde in [93] vorgestellt und basiert auf lokalen Extrema. Dies soll hier aufgegriffen werden. Entlang des Strahls der Richtung  $\mathbf{d} = \frac{\mathbf{x}-\mathbf{s}}{|\mathbf{x}-\mathbf{s}|}$  werden zunächst die richtungsgewichteten Gradienten ähnlich zu Gleichung 4.29 berechnet:

$$G_j = \begin{cases} 0 & ; \angle(\mathbf{I}'(\mathbf{x}_j), -\mathbf{d}) \geq 90^\circ \\ |\mathbf{I}'(\mathbf{x}_j)| \cdot \frac{\cos(\angle(\mathbf{I}'(\mathbf{x}_j), -\mathbf{d}))}{\cos \gamma} & ; \gamma < \angle(\mathbf{I}'(\mathbf{x}_j), -\mathbf{d}) < 90^\circ \\ |\mathbf{I}'(\mathbf{x}_j)| & ; \angle(\mathbf{I}'(\mathbf{x}_j), -\mathbf{d}) \leq \gamma \end{cases} \quad (4.30)$$

wobei  $\mathbf{x}_j = \mathbf{s} + j \frac{r_{\max}}{m} \mathbf{d}$  dem  $j$ -ten Punkt auf dem Strahl entspricht. Anschließend wird daraus die Menge  $\mathcal{M}$  berechnet, welche alle lokalen Maxima von  $G_j$  enthält:

$$j \frac{r_{\max}}{m} \in \mathcal{M} \iff G_j \geq G_{j-1} \wedge G_j > G_{j+1} \quad (4.31)$$

|                 | Logarithmische Gradientenstärke | Richtungsgewichtete log. Gradientenstärke | Lokale Gradientenstärkemaxima |
|-----------------|---------------------------------|---|-------------------------------|
| $\mathcal{T}_1$ | 0.56 mm, 0.13                   | 0.38 mm, 0.12, 33.80°                     | 0.37 mm, 0.12, 33.80°         |
| $\mathcal{T}_2$ | 0.58 mm, 0.13                   | 0.34 mm, 0.08, 36.10°                     | 0.34 mm, 0.08, 36.10°         |
| $\mathcal{T}_3$ | 0.60 mm, 0.13                   | 0.37 mm, 0.13, 32.66°                     | 0.36 mm, 0.13, 33.23°         |

Tabelle 4.1.: Die optimalen Parameter  $(\delta, \sigma)$  bzw.  $(\delta, \sigma, \gamma)$  der bildgradientenbasierten Kostenfunktionen wurden durch den Simplex-Optimierungs-Algorithmus für die verschiedenen Testmengen der Kreuzvalidierung bestimmt.

Die Gradientenkosten ergeben sich dann direkt aus dieser Menge zu

$$C_{\text{LokMaxGrad}}(\mathbf{x}) = \min_{\forall a \in \mathcal{M}} \left( 1 - e^{\left( \frac{|\mathbf{x} - \mathbf{s}| - a}{\psi} \right)^2} \right), \quad (4.32)$$

wobei der Gausskernel mit Standardabweichung  $\psi = 0.5 \text{ mm}$  eine örtliche Unsicherheit modelliert. Unabhängig von der Bildgradientenstärke sind die Kosten an den lokalen Maxima minimal (siehe Abbildung 4.8(c)).

Freie Parameter sind, wie bei der vorherigen Kostenfunktion auch, der Bildglättungsparameter  $\sigma$ , die systematische Abweichung  $\delta$  und der Winkel  $\gamma$ .

**Bestimmung der Parameter:** Anhand der gegebenen Referenzsegmentierungen werden nun die freien Parameter  $\sigma$ ,  $\delta$  und ggf.  $\gamma$  der beschriebenen bildgradientenbasierten Kostenfunktionen durch ein Optimierungsverfahren bestimmt. Aus jeder Referenzsegmentierung wird zunächst der Referenzradiusvektor  $\mathbf{r}^*$  bestimmt, welcher für jeden Strahl den Referenzradius enthält. Dies geschieht, indem ausgehend vom Saatpunkt  $\mathbf{s}$  entlang des Strahls iteriert und der Ort des Übergangs  $\mathbf{y}_i^*$  aus der Referenzsegmentierung detektiert wird:  $\mathbf{r}_i^* = |\mathbf{y}_i^* - \mathbf{s}|$ .

Für möglichst viele Strahlen sollte sich an diesen Stellen  $\mathbf{r}_i^*$  ein lokales Minimum der Kostenfunktion befinden und ansonsten möglichst wenig lokale Minima vorhanden sein. Das Vorliegen des globalen Minimums ist nicht zwingend erforderlich, weil zusammenhängende lokale Minima durch das Einfließen von Formwissen über benachbarte Strahlen hinweg ein globales Minimum der Gesamtkostenfunktion (siehe Gleichung 4.19) ergeben können. Darüber hinaus sind die später beschriebenen intensitätsbasierten Kostenfunktionen eher dazu geeignet, eine globale Einschränkung zu geben. Dennoch sollten die Kosten an den Referenzradien gering im Vergleich zu den umliegenden Kosten sein.

Die erste geforderte Eigenschaft, das Vorhandensein eines lokalen Kostenfunktionminimums am Rand der Referenzsegmentierung, kann durch eine lokale Erfolgsrate  $\nu \in [0, 1]$  beschrieben werden:

$$\nu = \frac{1}{n} \sum_{\forall i} \begin{cases} 1 & ; \min_{\forall a \in \mathcal{L}_i} |a - r_i^*| < \epsilon \\ 0 & ; \min_{\forall a \in \mathcal{L}_i} |a - r_i^*| \geq \epsilon \end{cases}, \quad (4.33)$$

wobei die Menge  $\mathcal{L}_i$  alle lokalen Minima der Kostenfunktion  $C_i$  entlang des Strahls  $i$  enthält. Die Fehlertoleranz wurde auf  $\epsilon = 0.5 \text{ mm}$  festgelegt.

Die zweite geforderte Eigenschaft, eine möglichst geringe Anzahl an lokaler Kostenfunktionsminima außerhalb des Referenzsegmentierungsrandes, kann durch die durchschnittliche Anzahl der lokalen



Kostenfunktionsminima beschrieben werden. Diese ist durch den Koeffizienten  $\eta \in \mathbb{R}_0^+$  gegeben:

$$\eta = \frac{1}{n} \sum_{\forall i} |\mathcal{L}_i| . \quad (4.34)$$

Die dritte geforderte Eigenschaft, vergleichsweise geringe Kosten am Referenzsegmentierungsrand, kann durch den Koeffizienten  $\xi \in \mathbb{R}_0^+$  ausgedrückt werden:

$$\xi = \frac{1}{n} \sum_{\forall i} \frac{C(\mathbf{s} + r_i^* \mathbf{d}_i)}{\frac{1}{m} \sum_{j=1}^m C(\mathbf{s} + j \frac{r_{\max}}{m} \mathbf{d}_i)} , \quad (4.35)$$

wobei der Nenner den durchschnittlichen Kosten entlang des Strahls  $i$  entspricht.

Für eine gegebene Menge  $\mathcal{T}$  von Intensitätsbildern mit dazugehörigen Saatpunkten und Referenzsegmentierungen können alle drei Koeffizienten als Funktion  $\nu_{\mathcal{T}}(\theta)$ ,  $\eta_{\mathcal{T}}(\theta)$  bzw.  $\xi_{\mathcal{T}}(\theta)$  der Energieparameter  $\theta = (\sigma, \delta)$  bzw.  $\theta = (\sigma, \delta, \gamma)$  aufgefasst werden, welche den durchschnittlichen Koeffizienten über diese Menge ausdrücken. Nun sollen die optimalen Parameter  $\theta_{\mathcal{T}}^*$  gefunden werden, also diejenigen, die  $\nu_{\mathcal{T}}(\theta)$  maximieren:

$$\theta_{\mathcal{T}}^* = \underset{\forall \theta}{\operatorname{argmax}} \nu_{\mathcal{T}}(\theta) . \quad (4.36)$$

Um ein lokales Optimierungsverfahren verwenden zu können, muss die Zielfunktion als *gutartig* angenommen werden. Sie muss also ein wohldefiniertes Maximum haben und in einem größeren Bereich um dieses Maximum herum konvex sein. Durch später beschriebene Experimente lässt sich begründet vermuten, dass diese Annahme tatsächlich zutrifft. Zum einen weist die erschöpfende Suche im Parameterraum in Abbildung 4.14 eine entsprechende Charakteristik auf und zum anderen werden erweisen sich die gefundenen optimalen Parameter für die unterschiedlichen Trainingsmengen der Kreuzvalidierung (siehe Abschnitt 5.1.1) als ähnlich (siehe Tabelle 4.1). Bei Funktionen mit mehreren lokalen Minima wäre dies i.d.R. nicht gegeben.

Das klassische Gradientenabstiegsverfahren oder ähnliche auf dem Gradienten der Zielfunktion basierende Optimierungsverfahren können jedoch nicht verwendet werden, weil eine analytische Berechnung dessen nicht praktikabel ist. Stattdessen wird hier der *Simplex-Algorithmus* [143] verwendet, welcher sich dadurch auszeichnet, dass er die partiellen Ableitungen der Zielfunktion nicht benötigt. Um zu verhindern, dass die Parameter  $\sigma$  und  $\gamma$  ihren Wertebereich verlassen und um gleichzeitig zu erreichen, dass ein potentiell größerer Wertebereich abgedeckt wird, findet eine Parametersubstitution statt:

$$\sigma = \sigma_0 \cdot 2^{\sigma'}, \quad (4.37)$$

$$\gamma = \alpha_0 \cdot 2^{\gamma'} . \quad (4.38)$$

So ist sichergestellt, dass  $\sigma$  und  $\gamma$  stets positiv sind. Eine Erhöhung (Verringerung) von  $\sigma'$  um 1 führt zu einer Verdopplung (Halbierung) von  $\sigma$ . Bei dem Parameter  $\delta$  ist dies nicht nötig, weil eine systematische Abweichung sowohl positiv, als auch negativ sein kann. Initial wurden die Parameter auf  $\sigma_0 = 0.5 \text{ mm}$  (ad hoc),  $\gamma_0 = 30^\circ$  (analog zu [88]) und  $\delta_0 = 0 \text{ mm}$  (kein systematischer Fehler) gesetzt. Der initiale Simplex wurde auf  $(1, 1, 1)$  gesetzt. Die Anzahl der Strahlen wird mit ca. 6500 auf einem sehr hohen Wert festgelegt. Im Zuge der Kreuzvalidierung wurde das Verfahren auf alle drei Trainingsmengen angewendet. Die erhaltenen Parameter sind in Tabelle 4.1 zu sehen. Erfreulicherweise sind diese für alle drei Mengen konsistent und es wird ersichtlich, dass in der Tat eine

|                         | Logarithmische<br>Gradientenstärke | Richtungsgewichtete<br>log. Gradientenstärke | Lokale Gradienten-<br>stärkemaxima |
|-------------------------|------------------------------------|--|------------------------------------|
| lok. Erfolgsrate $\nu$  | $0.64 \pm 0.11$                    | $0.66 \pm 0.08$                              | $0.65 \pm 0.09$                    |
| Anzahl Lok. Min. $\eta$ | $11.38 \pm 2.55$                   | $9.31 \pm 1.93$                              | $9.64 \pm 2.11$                    |
| Rel. Kosten $\xi$       | $0.24 \pm 0.17$                    | $0.34 \pm 0.15$                              | $0.52 \pm 0.15$                    |

Tabelle 4.2.: Durchschnitt und Standardabweichung der Koeffizienten der auf den bildgradientenbasierten Kostenfunktionen mit optimierten Parametern

systematische Abweichung zwischen der Referenzkontur und Orten mit lokalen Bildgradientenstärkemaxima vorliegt.

Anschließend wurden die Koeffizienten  $\nu$ ,  $\xi$  und  $\eta$  auf den Testmengen der Kreuzvalidierung (siehe Abschnitt 5.1.1) berechnet. Durchschnitt und Standardabweichung sind in Tabelle 4.2 eingetragen. Bei lokaler Erfolgsrate und Anzahl der lokalen Maxima erzielten die richtungsgewichtete logarithmische Gradientenstärke und die lokalen Gradientenstärkemaxima die besten Werte. Bei den relativen Kosten ist dies hingegen die logarithmische Gradientenstärke gefolgt von der richtungsgewichteten logarithmischen Gradientenstärke.

#### 4.1.6. Intensitätsbasierte Kostenfunktionen

Bildgradientenbasierte Kostenfunktionen sind zwar in der Lage, die Objektgrenze zu modellieren, sie können jedoch nichts über die Beschaffenheit im Inneren des Objektes aussagen. So kann es bspw. passieren, dass mehrere aneinander angrenzende Objekte als ein einzelnes Objekt von einem Segmentierungsalgorithmus identifiziert werden. Bei der Analyse des Inneren würde man diese fehlerhafte Segmentierung allerdings feststellen können. Daher ist es sinnvoll, schon von Anfang an die Beschaffenheit im Inneren des Lymphknotens in den Segmentierungsprozess einzubeziehen. Dies geschieht mittels intensitätsbasierter Kostenfunktionen. Im Folgenden werden drei Kostenfunktionen, davon eine Standardkostenfunktion, eine aus aktueller Literatur entnommene und eine in dieser Arbeit neu entwickelte Kostenfunktion, motiviert und beschrieben.

**Homogenität:** Wie bereits in Abschnitt 2.1.3 beschrieben, ist das Innere von gutartigen Lymphknoten ziemlich homogen. Die Grauwerte im Inneren sind um 80 HU näherungsweise gaussverteilt. Die Standardabweichung beträgt lediglich 20 HU. Diese Homogenität kann in eine intensitätsbasierte Kostenfunktion einfließen. Entlang eines Strahls wird im Inneren des Lymphknotens eine deutlich geringere Varianz erwartet als für Bereiche des Strahls, die die Lymphknotengrenze überschreiten. Dies kann für den  $j$ -ten Punkt  $\mathbf{p}_j = \mathbf{s} + j \frac{r_{\max}}{m} \mathbf{d}_i$  auf dem Strahl  $i$  durch eine Kostenfunktion der Form

$$C_{\text{Hom}}(\mathbf{p}_j) = \log \left( 1 + \frac{1}{j} \sum_{k=1}^j (G_{\sigma} \star I(\bar{\mathbf{p}}_k) - \mu_j)^2 \right) \quad (4.39)$$

beschrieben werden, wobei  $\bar{\mathbf{p}}_k = \mathbf{s} + (k \frac{r_{\max}}{m} - \delta) \mathbf{d}_i$  die korrigierte Version von  $\mathbf{p}_k$  und  $\mu_j = \frac{1}{j} \sum_{k=1}^j (I(G_{\sigma} \star \bar{\mathbf{p}}_k))$  die durchschnittliche Intensität ist. Der Term hängt also direkt von der Varianz der Grauwerte vom Saatpunkt bis zum Kandidaten ab (siehe Abbildung 4.10(a)). Die Faltung mit

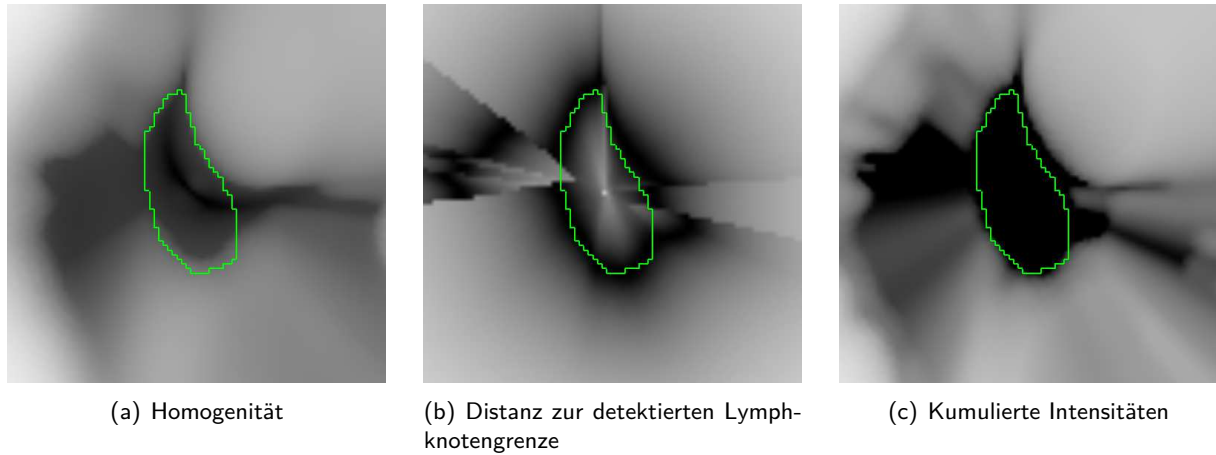


Abbildung 4.10.: Intensitätsbasierte Kostenfunktionen im Vergleich

dem Gausskernel  $G_\sigma$  erhöht die Toleranz gegenüber Rauschen und der Logarithmus sorgt dafür, dass sich die Kosten an allen Stellen in derselben Größenordnung befinden. Durch Letzteres ist sichergestellt, dass einzelne Bereiche im Inneren des Lymphknotens, welchen fälschlicherweise hohe Intensitätskosten zugeordnet wurden, durch die bildgradientenbasierten Kostenfunktionen oder durch die Übergangskosten dennoch korrekt segmentiert werden können.

Bei dieser Kostenfunktion beachte man allerdings, dass die Homogenitätsannahme im Inneren des Lymphknotens nur für gesunde Lymphknoten gilt. Lymphknoten mit zentralem nekrotischen Gewebe erfüllen diese Annahme hingegen nicht, weil das nekrotische Gewebe im CT etwas dunkler abgebildet wird (im Bereich um ca. 10 HU). Im Vergleich zum umliegenden Gewebe, wie z.B. Knochen oder kontrastverstärkten Gefäßen fällt diese Inhomogenität jedoch klein aus, d.h. dass im Inneren des Lymphknotens eine geringe Homogenität herrscht, diese aber dennoch höher als die Homogenität außerhalb des Lymphknotens ist. Freie Parameter sind der Bildglättungsparameter  $\sigma$  und die systematische Abweichung  $\delta$ .

**Distanz zur mutmaßlichen Lymphknotengrenze:** Wie bereits beschrieben, sind die Bildintensitäten im Inneren des Lymphknoten bekannt. Dieses Wissen kann zur Definition einer Kostenfunktion direkt herangezogen werden. Für jeden Strahl wird zunächst derjenige Index  $y$  identifiziert, welcher ausgehend vom  $l$ -ten Punkt als erstes ein vorgegebenes Intensitätsintervall  $[I_{\min}, I_{\max}]$  verlässt:

$$y = \min_{j \in \mathcal{I}} j, \quad j \in \mathcal{I} \iff j > l \wedge G_\sigma \star I(\bar{\mathbf{p}}_j) \notin [I_{\min}, I_{\max}] . \quad (4.40)$$

Der Punkt  $p_l$  ist dabei derjenige, dessen Intensität sich vom Saatpunkt ausgehend als erstes innerhalb des Intensitätsintervalls befindet (siehe Abbildung 4.11). Auf diese Art und Weise werden, unter der Annahme, dass der Saatpunkt sich im Inneren von etwaigem nekrotischen Gewebe befindet, auch Lymphknoten mit zentralem nekrotischen Gewebe unterstützt. Der so bestimmte Punkt  $\mathbf{p}_y$  kann nun als die, aufgrund der Bildintensitäten, wahrscheinlichste Lymphknotengrenze angesehen werden. Das Lymphknotenintensitätsintervall wurde auf [43, 119] festgelegt. Nach der Analyse in Abschnitt 2.1.3 entspricht dies einer Abdeckung von 95% aller im Inneren von nicht-nekrotischen Lymphknoten vorkommenden Intensitätswerte. Die geringeren Bildintensitäten von nekrotischem Gewebe werden

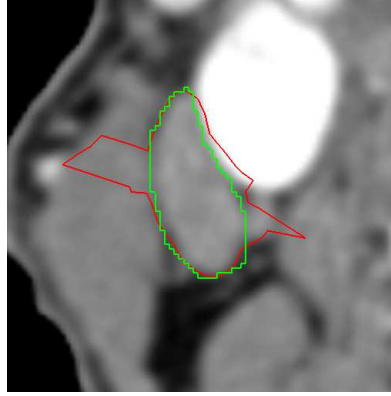


Abbildung 4.11.: Mutmaßliche Lymphknotenkontur aufgrund der Intensitätswerte entlang der Radialstrahlen (rot) im Vergleich zur Referenzsegmentierung (grün)

an dieser Stelle nicht mehr betrachtet, weil dieses zentral angenommen wurde und somit schon durch den Start der Detektion am  $l$ -ten Punkt abgedeckt ist.

Die Kosten aller Punkte des Strahls werden dann abhängig von dem Abstand zu diesem Punkt berechnet (siehe Abbildung 4.10(b)):

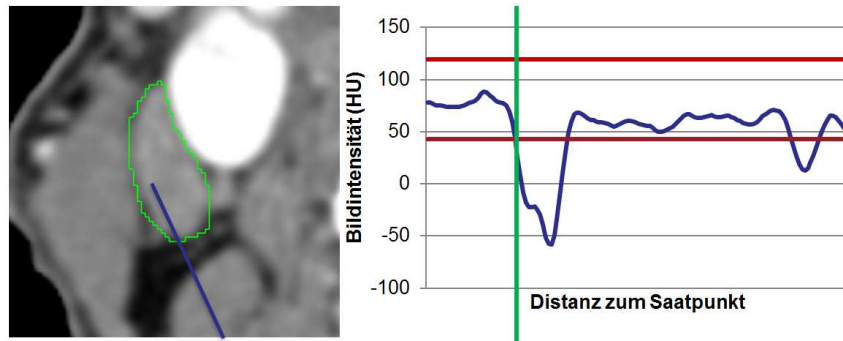
$$C_{\text{Dist}}(\mathbf{p}_j) = \log \left( 1 + \left( \frac{r_{\max}}{m} (j - y) \right)^2 \right). \quad (4.41)$$

Diese Kostenfunktion ist fast identisch mit der in [67] formulierten Kostenfunktion. Der Hauptunterschied besteht in der Bestimmung von  $y$ , welches dort auf eine deutlich einfachere Weise bestimmt wird: Es ist der erste Punkt entlang des Strahls, dessen Bildintensität um mehr als 50 HU von der des Saatpunkts abweicht. Freie Parameter sind, wie in der zuvor beschriebenen Kostenfunktion, der Bildglättungsparameter  $\sigma$  und die systematische Abweichung  $\delta$ .

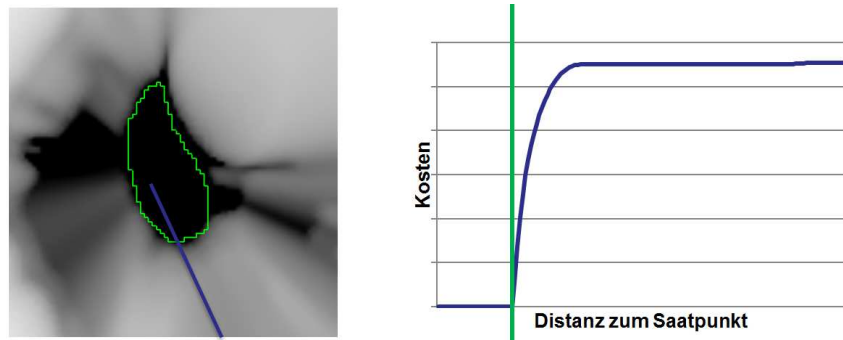
**Kumulierte Intensitäten:** Die gerade beschriebene Kostenfunktion wächst mit der Distanz zur mutmaßlichen Lymphknotengrenze. Dies hat den Nachteil, dass Objektgrenzen, welche z.B. aufgrund von geringem Kontrast zu Nachbarstrukturen nicht detektiert werden konnten oder durch Rauschen falsche Grenzen innerhalb des Lymphknotens detektiert werden, zu sehr hohen Kosten an der tatsächlichen Lymphknotengrenze führen können, die möglicherweise nicht von anderen Kostenfunktionen oder vom Formwissen ausgeglichen werden können. Daher ist es sinnvoll, die Kostenfunktion so zu erweitern, dass innerhalb des Lymphknotens keine Kosten entstehen und außerhalb die Kosten mit dem Abstand der Bildintensität zum Intervall  $[I_{\min}, I_{\max}]$  und der Distanz zur mutmaßlichen Lymphknotengrenze steigt. Beides kann ähnlich zu Gleichung 4.10 durch eine Kostenfunktion der Form

$$C_{\text{KumInt}}(\mathbf{p}_j) = \log \left( 1 + \sum_{k>l}^j \frac{r_{\max}}{m} \begin{cases} I(\bar{\mathbf{p}}_k) - I_{\max} & ; I(\bar{\mathbf{p}}_k) > I_{\max} \\ I_{\min} - I(\bar{\mathbf{p}}_k) & ; I(\bar{\mathbf{p}}_k) < I_{\min} \\ 0 & ; \text{sonst} \end{cases} \right) \quad (4.42)$$

erreicht werden. Abbildung 4.12 illustriert den Kostenverlauf beispielhaft entlang eines Strahls. Das dieser Kostenfunktion zugrunde liegende Prinzip hat sich bereits zur 2D-Segmentierung von Lymphknoten bewährt (vgl. Abschnitt 4.1.2). Das Intensitätsintervall wird wie bei der zuvor beschriebenen



(a) Grauwertprofil entlang des Strahls



(b) Kostenprofil entlang des Strahls

Abbildung 4.12.: Darstellung der kumulierten Intensitäten beispielhaft entlang eines Strahls (blau). Die vertikale grüne Linie entspricht der tatsächlichen Lymphknotengrenze, die horizontalen roten Linien markieren das bekannte Bildintensitätsintervall.

Kostenfunktion ebenfalls auf [43, 119] festgelegt und auch nekrotisches Gewebe wird auf dieselbe Art und Weise unterstützt: Das Kumulieren beginnt erst ab dem  $l$ -ten Punkt, welcher der erste entlang des Strahls ist, dessen Bildintensität sich innerhalb des vorgegebenen Intervalls befindet. Wie zuvor sind die einzigen beiden freien Parameter der Bildglättungsparameter  $\sigma$  und die systematische Abweichung  $\delta$ .

**Bestimmung der Parameter:** Wie bei den bildgradientenbasierten Kostenfunktionen hat die Wahl der freien Parameter einen entscheidenden Einfluss auf das Segmentierungsergebnis. Jedoch steht bei den intensitätsbasierten Kostenfunktionen weniger die genaue Lokalisierung der Lymphknotengrenze im Vordergrund, sondern das globale Ausschließen von Bereichen, welche aufgrund der Intensitätswerte im Inneren nicht zu dem zu segmentierenden Lymphknoten gehören können. Der zu optimierende Koeffizient – die globale Erfolgsrate  $\nu'$  – bezieht sich daher auf das globale Minimum der Kostenfunktion:

$$\nu' = \frac{1}{n} \sum_{\forall i} \begin{cases} 1 & ; \left| \left( \argmin_{\forall j} E_i(\mathbf{p}_j) \right) \frac{r_{\max}}{m} - r_i^* \right| < \epsilon \\ 0 & ; \left| \left( \argmin_{\forall j} E_i(\mathbf{p}_j) \right) \frac{r_{\max}}{m} - r_i^* \right| \geq \epsilon \end{cases} . \quad (4.43)$$

|                 | Homogenität   | Distanz zur mutmaßlichen Lymphknotengrenze | Kumulierte Intensitäten |
|-----------------|---------------|--|-------------------------|
| $\mathcal{T}_1$ | 0.23 mm, 2.19 | 0.67 mm, 0.24                              | 0.81 mm, 0.24           |
| $\mathcal{T}_2$ | 0.31 mm, 2.16 | 0.60 mm, 0.22                              | 0.78 mm, 0.22           |
| $\mathcal{T}_3$ | 0.21 mm, 2.20 | 0.64 mm, 0.26                              | 0.81 mm, 0.24           |

Tabelle 4.3.: Die optimalen Parameter  $(\delta, \sigma)$  der bildintensitätsbasierten Kostenfunktionen wurden durch den Simplex-Optimierungs-Algorithmus für die verschiedenen Testmengen der Kreuzvalidierung bestimmt.

|                               | Homogenität     | Distanz zur mutmaßlichen Lymphknotengrenze | Kumulierte Intensitäten |
|-------------------------------|-----------------|--|-------------------------|
| glob. Erfolgsrate $\nu'$      | $0.06 \pm 0.07$ | $0.54 \pm 0.22$                            | $0.43 \pm 0.24$         |
| Rel. Kosten $\xi$             | $0.35 \pm 0.15$ | $0.26 \pm 0.22$                            | $0.08 \pm 0.07$         |
| Distanz zu glob. Min. $\zeta$ | $3.85 \pm 2.55$ | $2.77 \pm 2.55$                            | $2.63 \pm 2.58$         |

Tabelle 4.4.: Durchschnitt und Standardabweichung der Koeffizienten für die bildintensitätsbasierten Kostenfunktionen mit optimierten Parametern

Dies steht im Gegensatz zur lokalen Erfolgsrate (vgl. Gleichung 4.33), welche bei den bildgradientenbasierten Kostenfunktionen maximiert wurde. Zusätzlich wurde die Fehlertoleranz mit  $\epsilon = 1 \text{ mm}$  etwas höher gesetzt. Der Optimierungsprozess ist ansonsten derselbe und wird wie zuvor auch schon im Zuge der Kreuzvalidierung auf alle drei Trainingsmengen (siehe Abschnitt 5.1.1) unabhängig angewendet. Es kommen ein Simplex-Optimierungsverfahren und die Parametersubstitution des Bildglättungsparameters  $\sigma$  zum Einsatz. Die optimalen Parameter sind in Tabelle 4.3 vermerkt. Auch hier sind die erhaltenen Werte für alle drei Trainingsmengen konsistent.

Auf den Testmengen wurden auch für diese Kosten aussagekräftige Koeffizienten berechnet. Neben der globalen Erfolgsrate  $\nu'$  und der in Gleichung 4.35 definierten relativen Kosten  $\xi$  ist dies auch der durchschnittliche Abstand der Referenzkontur zum globalen Minimum, welcher durch

$$\zeta = \frac{1}{n} \sum_{\forall i} \left| \left( \underset{\forall j}{\operatorname{argmin}} E_i(\mathbf{p}_j) \right) \frac{r_{\max}}{m} - r_i^* \right| \quad (4.44)$$

in mm gegeben ist. Der letztere Koeffizient ist ein Maß für die Güte der globalen Lokalisation der Kostenfunktion. Da dies im Vordergrund steht, wurde auf die Berechnung der durchschnittlichen Anzahl an lokalen Minima  $\eta$  verzichtet. Die Ergebnisse sind in Tabelle 4.4 eingetragen.

Aufgrund dieser Werte können für die Kostenfunktionen folgende Aussagen getroffen werden. Durch die geringe Erfolgsrate, hohen relativen Kosten an der tatsächlichen Lymphknotengrenze und der hohen Distanz zum globalen Minimum scheint die Homogenität im Inneren als exklusive Kostenfunktion nicht geeignet zu sein. Die anderen beiden Kostenfunktionen zeigen ähnliche Charakteristiken, wobei die kumulierten Intensitäten durch sehr geringe relative Kosten an der tatsächlichen Lymphknotengrenze herausstechen und leicht geringere Distanz von dem Referenzsegmentierungsrand zum globalen Kostenminimum aufweisen.

#### 4.1.7. Parameterbestimmung

Die in dem vergangenen Abschnitt vorgestellten und untersuchten Kostenfunktionen sollen nun in einem Radialstrahl-basierten Segmentierungsverfahren zum Einsatz kommen. Ziel dieses Abschnittes ist es, die beste Kostenfunktionskombination zu identifizieren und geeignete Gewichte der Kostenfunktionen untereinander und die Parameter der formerhaltenden Kosten zu bestimmen. Als Gütekriterium dient der Index  $\mathcal{G} \in [0, 100]$ , welcher ein Maß für die Ähnlichkeit zwischen einer durch einen Algorithmus erzeugten Segmentierung und einer Referenzsegmentierung ist (100 entspricht einer perfekten Überlappung, 75 entspricht der Qualität von manueller Segmentierung). Eine Definition dieses Maßes ist in Abschnitt 5.1.1 zu finden.

**Einzelne Kostenfunktionen:** In einem ersten Schritt wird die Eignung der Kostenfunktionen zur Radialstrahl-basierten Segmentierung von Lymphknoten jeweils für sich geprüft. Da unterschiedliche Optimierungsverfahren mit unterschiedlichen Kostenfunktionen besser/schlechter als andere zurechtkommen können, kommen alle drei in Abschnitt 4.1.4 beschriebenen Optimierungsverfahren (Gradientenabstieg (GA), Belief Propagation (BP) und Graph Cut (GC)) mit zugehöriger Formerhaltung zum Einsatz.

Um den Einfluss der Kostenfunktion auf die Segmentierungsgüte möglichst unabhängig von allen anderen Faktoren feststellen zu können, wurden möglichst viele Parameter konstant gesetzt. Für GA und BP ist einzig und allein der Formerhaltungskostenparameter  $\lambda$  variabel, weil für diesen Parameter offensichtlich für unterschiedliche Kostenfunktionen unterschiedliche Werte zum besten Segmentierungsergebnis führen. Bei der mit GC verwendeten Formerhaltung existiert ein solcher Parameter nicht.

Die Anzahl der Strahlen wurde ad hoc auf etwas über 2000, die Anzahl der Punkte auf einem Strahl auf  $m = 120$  und der maximale Radius auf  $r_{\max} = 30 \text{ mm}$  festgelegt. Dies hat zur Folge, dass Lymphknoten bis zu einem großen Durchmesser von  $60 \text{ mm}$  unterstützt werden und die maximale Segmentierungsgenauigkeit bei  $\frac{r_{\max}}{m} = 0.25 \text{ mm}$  liegt, was sich größenordnungsmäßig im Bereich der üblichen Auflösung innerhalb von CT-Schichten befindet.

Der Exponent der Formerhaltungskosten des GA- und des BP-Verfahrens wurde auf  $\alpha = 2$  gesetzt. Der Formerhaltungsparameter  $\Delta$  des GC-Verfahrens wurde folgendermaßen bestimmt: Aus dem Referenzradienvektor  $\mathbf{r}^*$  des  $k$ -ten Lymphknotens wurde die maximale Differenz benachbarter Strahlen bestimmt:

$$\Delta_k = \max_{\forall i, j \in \mathcal{N}(i)} \left[ m \frac{|r_i^* - r_j^*|}{r_{\max}} \right]. \quad (4.45)$$

$\Delta =$  wurde dann so gewählt, dass 95% aller maximalen Differenzen  $\Delta_k$  abgedeckt werden können.

Für GA und BP soll nun der optimale Formerhaltungskostenparameter  $\lambda$  so bestimmt werden, dass die mittlere Güte  $\bar{\mathcal{G}}_{\mathcal{T}}$  einer Trainingsmenge  $\mathcal{T}$  maximiert wird.  $\bar{\mathcal{G}}_{\mathcal{T}}$  wird dabei folgendermaßen bestimmt: Der Segmentierungsalgorithmus wird auf jedes Element der Menge  $\mathcal{T}$  angewendet, wobei der Saatpunkt jeweils in den Schwerpunkt der Referenzsegmentierung gelegt wird. Anschließend wird das Gütemaß  $\mathcal{G}$  zwischen dem Segmentierungsergebnis und der Referenzsegmentierung berechnet.  $\bar{\mathcal{G}}_{\mathcal{T}}$  ergibt sich dann zu dem Mittelwert der Gütemaße  $\mathcal{G}$  und kann als Funktion des Formerhaltungsparameters  $\lambda$  aufgefasst werden. Die Bestimmung des optimalen Segmentierungsparameters  $\lambda^*$  kann

| GA              | Abs. Grad. | Richt. Grad. | Lok. Max. | Homogenität | Distanz | Kum. Int. |
|-----------------|------------|--------------|-----------|-------------|---------|-----------|
| $\mathcal{T}_1$ | 50         | 100          | 50        | 100         | 2       | 0.09      |
| $\mathcal{T}_2$ | 5.55       | 50           | 50        | 100         | 16.66   | 0.15      |
| $\mathcal{T}_3$ | 33.33      | 7.69         | 50        | 100         | 7.14    | 0.12      |

Tabelle 4.5.: Die optimalen Formerhaltungskosten  $\lambda^*$  der einzelnen Trainingsmengen der Kreuzvalidierung für die verschiedenen einzelnen Kostenfunktionen unter Verwendung des Gradientenabstiegsverfahrens

| BP              | Abs. Grad. | Richt. Grad. | Lok. Max. | Homogenität | Distanz | Kum. Int. |
|-----------------|------------|--------------|-----------|-------------|---------|-----------|
| $\mathcal{T}_1$ | 2.56       | 1.45         | 0.08      | 0.01        | 1.85    | 0.01      |
| $\mathcal{T}_2$ | 1.96       | 1.12         | 0.07      | 0.01        | 2.04    | 0.01      |
| $\mathcal{T}_3$ | 2.00       | 1.11         | 0.09      | 0.01        | 2.04    | 0.81      |

Tabelle 4.6.: Die optimalen Formerhaltungskosten  $\lambda^*$  der einzelnen Trainingsmengen der Kreuzvalidierung für die verschiedenen einzelnen Kostenfunktionen unter Verwendung des Belief-Propagation-Verfahrens

nun als Funktionsmaximierungsproblem aufgefasst werden:

$$\lambda^* = \underset{\forall \lambda}{\operatorname{argmax}} \bar{\mathcal{G}}_{\mathcal{T}}(\lambda) . \quad (4.46)$$

Unter der Annahme, dass die Zielfunktion gutartig ist, kann, wie schon zur Bestimmung der Kostenfunktionsparameter, der Simplex-Algorithmus verwendet werden. Um einen großen Bereich abdecken zu können und  $\lambda > 0$  garantieren zu können, wurde die Parametersubstitution  $\lambda = \lambda_0 \cdot 2^{\lambda'}$  angewendet. Der Startwert wurde auf  $\lambda_0 = 1$  festgelegt, was einer Radienabweichung benachbarter Strahlen  $i, j$  von 1 mm die Formerhaltungskosten  $C_{i,j} = 1$  zuordnet. Die initiale Größe des Simplex wurde auf 1 festgelegt. Die Optimierung wurde für die drei Trainingsmengen separat für alle zuvor beschriebenen Kostenfunktionen durchgeführt. Der optimale Parameter wurde für alle Kostenfunktionen in Tabelle 4.5 für das Gradientenabstiegsverfahren bzw. in Tabelle 4.6 für das Belief-Propagation-Verfahren eingetragen. Anschließend wurden die unterschiedlichen Optimierungsverfahren mit den gefundenen Parametern zur Segmentierung aller Elemente der Testmengen angewendet und für jedes Element die Segmentierungsgüte  $\mathcal{G}$  bestimmt. Statistiken dazu sind in Tabelle 4.7 zu finden.

Bei den bildgradientenbasierten Verfahren eignet sich die richtungsgewichtete logarithmische Gradientenstärke am besten. Beide globale Optimierungsverfahren (GC und BP) erzielen deutlich bessere Werte als das Gradientenabstiegsverfahren, wobei GC mit 69.69 eine bessere durchschnittliche Segmentierungsgüte als BP mit 57.88 erreichen kann. Dies muss allerdings nicht notwendigerweise an dem Optimierungsverfahren liegen, sondern kann auch an den unterschiedlichen Arten der Formerhaltung liegen, deren Parameter bisher nur ad hoc festgelegt wurden und damit möglicherweise im Falle von GC geeigneter sein könnten. Der Grund, weshalb GA so schlechte Werte liefert, ist vermutlich auf eine schlechte Initialisierung und die daraus resultierende Konvergenz zu einem falschen lokalen Minimum zurückzuführen. Die Initialisierung durch globale Minima der Kostenfunktion für jeden Strahl ist damit bei den bildgradientenbasierten Kostenfunktionen ungeeignet. Dies ist auch nicht weiter verwunderlich, weil starke Bildgradientenstärken an vielen Stellen im umliegenden Gewebe vorkommen. Globale Optimierungsverfahren benötigen keine Initialisierung und sind daher nicht betroffen.



|                 | GA    | BP    | GC    |
|-----------------|-------|-------|-------|
| Abs. Grad.      | 5.38  | 33.74 | 27.21 |
| Richt. Grad.    | 11.70 | 57.88 | 69.69 |
| Lok. Max. Grad. | 3.21  | 28.47 | 49.57 |
| Homogenität     | 15.21 | 15.33 | 50.44 |
| Distanz         | 33.74 | 55.31 | 15.99 |
| Kum. Int.       | 33.71 | 22.77 | 45.06 |

Tabelle 4.7.: Durchschnittliche Segmentierungsgüte  $\bar{G}$  für die verschiedenen einzelnen Kostenfunktionen der unterschiedlichen Segmentierungsverfahren

Die bildintensitätsbasierten Kostenfunktionen eignen sich in der Regel schlechter als die bildgradientenbasierten Kostenfunktionen. GA bildet allerdings eine Ausnahme dazu, was zeigt, dass bildintensitätsbasierte Kostenfunktionen besser in der Lage sind, die ungefähre Größe zu lokalisieren. Dies entspricht der Konstruktionsabsicht. Unklar bleibt jedoch, weshalb BP bei der Distanz zur Lymphknotengrenze und GC bei der Homogenität eine herausstechend hohe Segmentierungsgüte erzielt.

Mit einer maximal erzielten durchschnittlichen Segmentierungsgüte von 69.69 für die richtungsgegewichteten logarithmischen Gradientenstärke und GC, sind die erhaltenen Segmentierungsergebnisse jedoch noch nicht zufriedenstellend. Im Vergleich zu manueller Segmentierung ist dieses Verfahren immer noch um 7% schlechter. Daraus kann geschlossen werden, dass weder die Betrachtung von Diskontinuität an der Objektgrenze, noch die Beschaffenheit im Inneren für sich ausreichend ist, um Lymphknoten mit hoher Genauigkeit segmentieren zu können.

**Kombination der Kostenfunktionen:** Da weder bildgradientenbasierte noch intensitätsbasierte Kostenfunktionen für sich alleine zur Segmentierung von Lymphknoten geeignet sind, werden diese nun additiv miteinander kombiniert. Die Bildenergie aus Gleichung 4.15 ergibt sich nun zu

$$E_{\text{Bild}}(\mathbf{l}) = \sum_{\forall i} \rho C_{i_{\text{Grad}}}(l_i) + C_{i_{\text{Int}}}(l_i) . \quad (4.47)$$

Dabei wurde ein neuer Parameter  $\rho$  eingeführt, welcher den relativen Einfluss der bildgradientenbasierten Kosten zu den intensitätsbasierten Kosten auf die Gesamtkosten regelt. Die optimalen Parameter  $\lambda^*$  und  $\rho^*$  wurden analog zur Bestimmung der optimalen Formerhaltungskosten  $\lambda^*$  im vorangegangenen Abschnitt durch Maximierung der durchschnittlichen Segmentierungsgüte für die unterschiedlichen Segmentierungsverfahren bestimmt. Bei GC muss lediglich der Parameter  $\rho^*$  gefunden werden, weil  $\lambda^*$  nicht existiert. Da eine geringe Eignung einer Kostenfunktion für sich alleine nicht notwendigerweise auch eine geringe Eignung in Kombination mit einer anderen Kostenfunktion nach sich ziehen muss, kann keine der vorgestellten Kostenfunktionen a priori ausgeschlossen werden. Daher wurde die Parameteroptimierung auf den Trainingsmengen und die anschließende Messung der Segmentierungsgüte auf den Testmengen für alle 9 möglichen Kombinationen aus bildgradientenbasierter und intensitätsbasierter Kostenfunktionen durchgeführt.

Die optimalen Parameter sind für die unterschiedlichen Optimierungsverfahren in Tabelle 4.8, 4.9 und 4.10 eingetragen. Für die verschiedenen Trainingsmengen befinden sich die durch die Simplex-Optimierung erzielten Werte in den allermeisten Fällen in derselben Größenordnung. Diese Konsistenz ist einerseits ein Anzeichen für die Validität des Parameterbestimmungsverfahrens und andererseits

#### 4. Methoden

| Gradientenabstieg |             | Homogenität |      |      | Distanz |      |       | Kum. Int. |       |       |
|-------------------|-------------|-------------|------|------|---------|------|-------|-----------|-------|-------|
| Abs. Grad.        | $\lambda^*$ | 2.22        | 3.33 | 2.00 | 2.70    | 4.76 | 1.25  | 0.28      | 0.24  | 0.27  |
|                   | $\rho^*$    | 3.22        | 3.32 | 3.30 | 1.95    | 1.94 | 1.77  | 0.04      | 0.003 | 0.002 |
| Richt. Grad.      | $\lambda^*$ | 7.14        | 7.69 | 5.26 | 10.0    | 10.0 | 10.0  | 0.28      | 0.27  | 0.32  |
|                   | $\rho^*$    | 2.62        | 2.66 | 2.45 | 1.96    | 1.96 | 1.96  | 0.06      | 0.07  | 0.08  |
| Lok. Max. Grad.   | $\lambda^*$ | 1.01        | 0.58 | 0.75 | 50.0    | 20.0 | 50.0  | 100       | 0.20  | 4.00  |
|                   | $\rho^*$    | 5.29        | 6.57 | 3.12 | 13.58   | 8.23 | 11.60 | 89.08     | 0.12  | 2.00  |

Tabelle 4.8.: Die optimalen Parameter  $\lambda^*$ ,  $\rho^*$  der gradientenbasierten gegenüber der intensitätsbasierten Kostenfunktion für das auf Gradientenabstieg basierende Segmentierungsverfahren. In jedem Feld sind die aus den drei disjunkten Trainingsmengen  $\mathcal{T}_1$ ,  $\mathcal{T}_2$  und  $\mathcal{T}_3$  der Kreuzvalidierung durch Optimierung bestimmten Werte eingetragen.

| Belief-Propagation |             | Homogenität |       |       | Distanz |      |       | Kum. Int. |      |      |
|--------------------|-------------|-------------|-------|-------|---------|------|-------|-----------|------|------|
| Abs. Grad.         | $\lambda^*$ | 0.72        | 9.09  | 12.5  | 3.03    | 5.26 | 4.35  | 3.33      | 1.49 | 4.00 |
|                    | $\rho^*$    | 4.05        | 5.66  | 6.92  | 2.67    | 2.86 | 1.97  | 2.68      | 2.38 | 3.45 |
| Richt. Grad.       | $\lambda^*$ | 10.0        | 7.69  | 12.5  | 9.09    | 4.76 | 7.69  | 7.69      | 6.25 | 3.57 |
|                    | $\rho^*$    | 5.72        | 5.32  | 6.29  | 5.19    | 2.93 | 4.76  | 5.55      | 5.42 | 3.34 |
| Lok. Max. Grad.    | $\lambda^*$ | 0.13        | 0.49  | 0.45  | 7.69    | 2.94 | 8.33  | 0.10      | 0.64 | 0.57 |
|                    | $\rho^*$    | 18.02       | 48.16 | 42.38 | 22.84   | 8.32 | 24.96 | 47.73     | 2.65 | 2.29 |

Tabelle 4.9.: Die optimalen Parameter  $\lambda^*$ ,  $\rho^*$  der gradientenbasierten gegenüber der intensitätsbasierten Kostenfunktion für das auf Belief-Propagation basierende Segmentierungsverfahren. In jedem Feld sind die aus den drei disjunkten Trainingsmengen  $\mathcal{T}_1$ ,  $\mathcal{T}_2$  und  $\mathcal{T}_3$  der Kreuzvalidierung durch Optimierung bestimmten Werte eingetragen.

| Graph-Cut  |          | Homogenität |      |      | Distanz |       |       | Kum. Int. |      |      |
|------------|----------|-------------|------|------|---------|-------|-------|-----------|------|------|
| Abs. Grad. | $\rho^*$ | 5.60        | 5.55 | 5.65 | 2.99    | 3.88  | 2.09  | 3.32      | 3.35 | 1.92 |
|            | $\rho^*$ | 3.32        | 5.63 | 5.73 | 8.02    | 7.85  | 6.91  | 3.67      | 3.37 | 1.92 |
| Lok. Max.  | $\rho^*$ | 9999        | 9999 | 9999 | 26.84   | 27.57 | 23.82 | 9999      | 9999 | 9999 |

Tabelle 4.10.: Das optimale Gewicht  $\rho^*$  der gradientenbasierten gegenüber der intensitätsbasierten Kostenfunktion für das auf Graph-Cut basierende Segmentierungsverfahren. In jedem Feld sind die aus den drei disjunkten Trainingsmengen  $\mathcal{T}_1$ ,  $\mathcal{T}_2$  und  $\mathcal{T}_3$  der Kreuzvalidierung durch Optimierung bestimmten Werte eingetragen.

|              | Homogenität |       |       | Distanz |       |       | Kum. Int. |       |       |
|--------------|-------------|-------|-------|---------|-------|-------|-----------|-------|-------|
|              | GA          | BP    | GC    | GA      | BP    | GC    | GA        | BP    | GC    |
| Abs. Grad.   | 28.50       | 51.35 | 55.62 | 35.29   | 59.29 | 50.20 | 44.37     | 66.68 | 65.20 |
| Richt. Grad. | 40.53       | 70.87 | 71.35 | 37.83   | 63.82 | 69.79 | 46.66     | 70.05 | 72.66 |
| Lok. Max.    | 16.51       | 22.99 | 49.52 | 36.11   | 60.19 | 71.13 | 31.16     | 29.14 | 49.55 |

Tabelle 4.11.: Durchschnittliche Segmentierungsgüte  $\bar{g}$  für die verschiedenen Kostenfunktionskombinationen der unterschiedlichen Segmentierungsverfahren

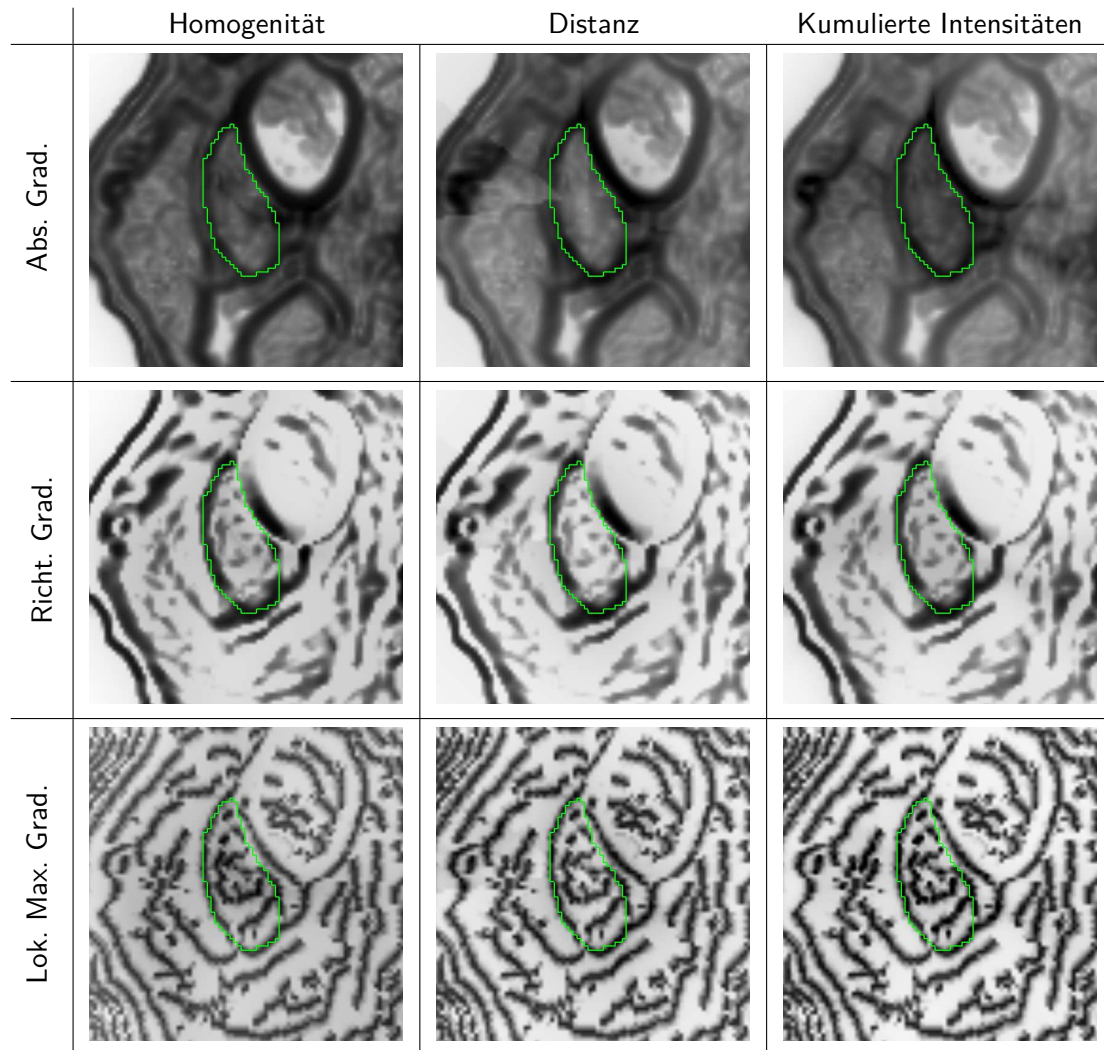


Abbildung 4.13.: Kombination bildgradientenbasierter und intensitätsbasierter Kostenfunktionen im Vergleich. Die bildgradientenbasierten Kostenfunktionen dominieren. Schon die geringfügige Erhöhung der Kosten außerhalb der Lymphknotengrenzen durch die intensitätsbasierten Kosten ist ausreichend, um die Segmentierungsgüte zu erhöhen.

für die Vergleichbarkeit der einzelnen Trainingsmengen. Ausnahmen bilden teilweise die Werte für die lokalen Maxima der Gradientenstärke. Ein Grund dafür könnte sein, dass die absolute Bildgradientenstärke bei diesem Verfahren überhaupt keinen Einfluss auf das Segmentierungsergebnis hat und so das Verfahren anfällig für zufällige lokale Gradientenstärkemaxima ist.

Wie in Abbildung 4.13 für BP zu sehen ist, dominiert der Einfluss der bildgradientenbasierten Kostenfunktionen den der intensitätsbasierten Kostenfunktionen. Wie zu erwarten, ist der ermittelte optimale Einfluss der bildgradientenbasierten Kostenfunktionen bei Verwendung des lokalen Verfahrens GA deutlich geringer als bei den globalen Verfahren BP und GC. Bei dem optimalen Formerhaltungsparameter  $\lambda^*$  zeichnet sich im Vergleich der verschiedenen Optimierungsverfahren kein Trend ab.

Die erzielte durchschnittliche Segmentierungsgüte, welche auf den Testmengen gemessen wurde, ist für die verschiedenen Kombinationsmöglichkeiten der Kostenfunktionen und für die verschiedenen Optimierungsverfahren in Tabelle 4.11 eingetragen. Die Kombination von bildgradientenbasierten und intensitätsbasierten Kostenfunktionen stellt einen deutlichen Mehrwert dar. Dennoch liefern die globalen Verfahren BP und GC nach wie vor erheblich bessere Ergebnisse als das lokale GA-Verfahren, wobei GC etwas bessere Werte als BP erzielt. Dieser geringe Unterschied kann möglicherweise, wie bereits erläutert, aber auch durch das unterschiedliche Formwissen bzw. mögliche bisher ad hoc festgelegte nicht-optimale Parameter zustande kommen.

Die besten Ergebnisse werden durch die Kombination der richtungsgewichteten logarithmischen Gradientenstärke mit den kumulierten Intensitäten für GC und der Homogenität für BP erzielt. Überraschenderweise erzielt das GC-Verfahren auch bei der Kombination der lokalen Gradientenstärkemaxima mit der Distanz zur mutmaßlichen Lymphknotengrenze eine sehr große durchschnittliche Segmentierungsgüte, obwohl die von den einzelnen Kostenfunktionsteilen erzielten Ergebnisse (siehe Tabelle 4.7) ganz und gar nicht vielversprechend aussehen.

**Anzahl der Radialstrahlen:** Die Kombination aus richtungsgewichteter logarithmischer Bildgradientenstärke und kumulierten Intensitäten bzw. Homogenität hat sich für das Gradientenabstiegs- und Graph-Cut- bzw. für das Belief-Propagation-Verfahren als am erfolgreichsten herausgestellt. Dabei wurden allerdings einige Parameter, welche sicherlich einen Einfluss auf das Segmentierungsergebnis haben, ad hoc festgelegt. Dies sind zum einen die Parameter der Formerhaltung ( $\alpha$ ,  $\lambda$  bzw.  $\Delta$ ) und zum anderen die Anzahl der Strahlen  $n$ . Diese Parameter beeinflussen sich gegenseitig. Je mehr Strahlen verwendet werden, desto näher sind die Radian benachbarter Strahlen und damit desto stärker muss die Formerhaltung sein. Dieser Zusammenhang soll nun für die bei GA und BP verwendete Formerhaltung näher untersucht werden. Für GC findet diese Analyse nicht statt, weil eine andere Formerhaltung zur Anwendung kommt, die den Parameter  $\lambda$  nicht enthält.

Sei  $\lambda_n^*$  der optimale Formerhaltungsparameter für  $n$  Strahlen. Der Einfachheit halber soll  $n$  hier die Anzahl der Strahlen in der Ebene angeben. Dadurch kann der Winkel  $\phi$  zwischen zwei Strahlen durch  $\phi = \frac{360^\circ}{n}$  berechnet werden. Die tatsächliche Anzahl der Strahlen auf der Kugeloberfläche ist proportional zu  $n^2$ . Weiterhin sei  $E_{Bild}$  die dazugehörige Bildenergie, welche durch Gleichung 4.15 berechnet wurde. Es wird angenommen, dass die Radiendifferenz benachbarter Strahlen an allen Stellen durch den konstanten Term  $d \frac{360^\circ}{n}$  näherungsweise gegeben ist.  $d > 0$  entspricht hierbei der Radiendifferenz pro Winkel. Die Formerhaltung aus Gleichung 4.17 vereinfacht sich unter dieser

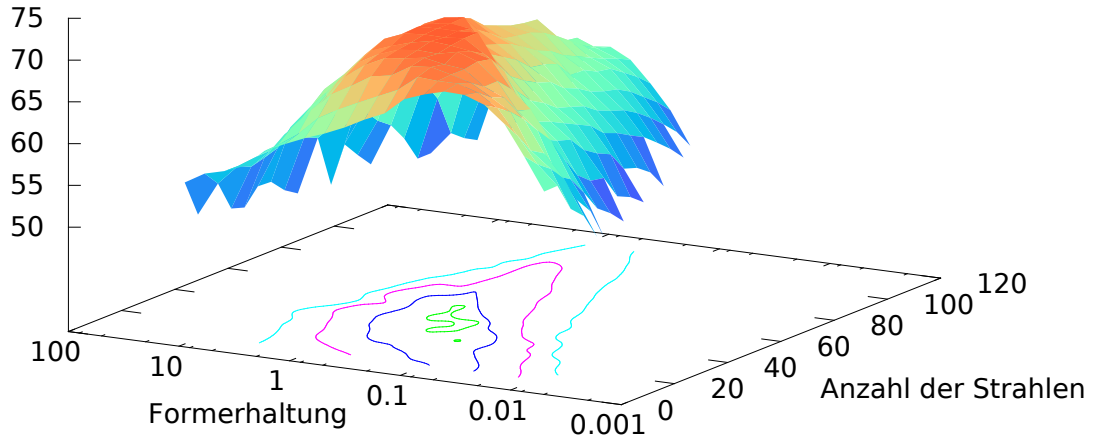


Abbildung 4.14.: Segmentierungsgüte des BP-Verfahrens für unterschiedliche Anzahl der Strahlen  $n$  und Formerhaltungskosten  $\lambda$

Annahme zu:

$$E_{\text{Form}} \approx n\lambda \left( d \frac{360^\circ}{n} \right)^\alpha. \quad (4.48)$$

Wird nun die Anzahl der Strahlen auf  $n'$  verändert, hat das Folgen sowohl für den Bildenergieterm als auch für den Formenergieterm. Da die Kosten an nahe beieinanderliegenden Orten ähnlich sind, ändern sich die Kosten der einzelnen Strahlen für ähnliche Strahlenanzahlen  $n$  und  $n'$  kaum. Allerdings ändert sich die Anzahl der Summanden in Gleichung 4.15 und der Bildenergieterm kann durch

$$E'_{\text{Bild}} \approx \frac{n'}{n} E_{\text{Bild}} \quad (4.49)$$

angenähert werden. Beim Formenergieterm kommen hingegen zwei gegenläufige Effekte zu tragen. Steigt die Anzahl der Strahlen, steigt auch die Anzahl der Summanden in Gleichung 4.17. Gleichzeitig sinkt allerdings die Radiendifferenz in Gleichung 4.16. Die Formerhaltungskosten  $\lambda$  müssen also so angepasst werden, dass sich das Verhältnis aus Bild- und Formenergie nicht ändert:

$$\frac{E_{\text{Bild}}}{E_{\text{Form}}} = \frac{E'_{\text{Bild}}}{E'_{\text{Form}}}. \quad (4.50)$$

Mit den gerade beschriebenen Näherungen ergibt sich daraus

$$\frac{E_{\text{Bild}}}{n\lambda \left( d \frac{360^\circ}{n} \right)^\alpha} = \frac{\frac{n'}{n} E_{\text{Bild}}}{n'\lambda' \left( d \frac{360^\circ}{n'} \right)^\alpha}. \quad (4.51)$$

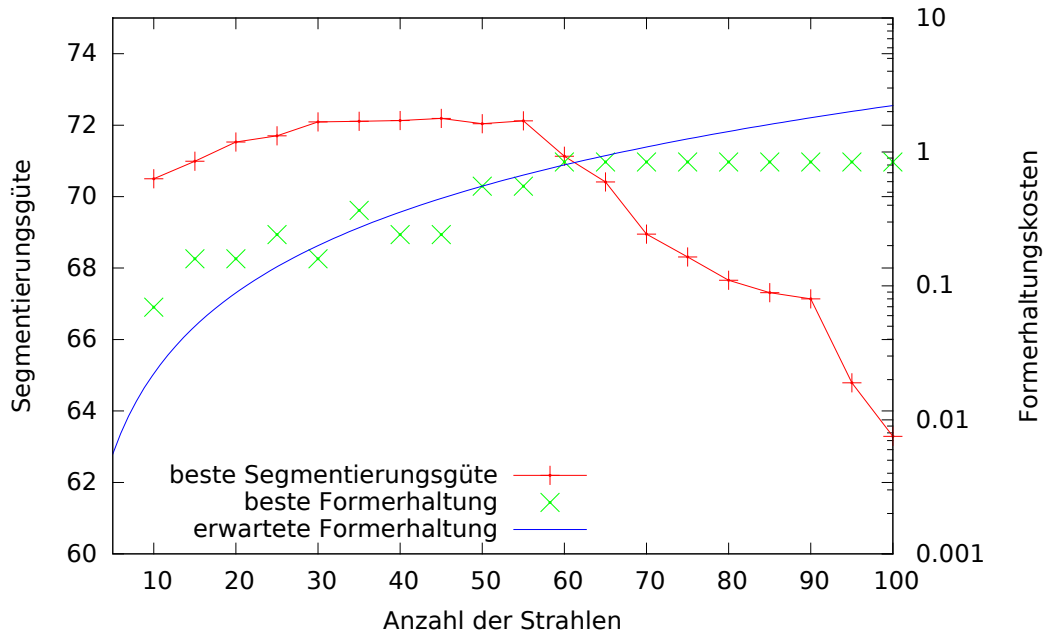


Abbildung 4.15.: Einfluss der Strahlenanzahl auf die Segmentierungsgüte bei der bestmöglichen Formerhaltung und gemessene und berechnete optimale Formerhaltungskosten  $\lambda^*$

Durch Umformung ergeben sich Formerhaltungskosten von

$$\lambda' = \lambda \left( \frac{n'}{n} \right)^\alpha. \quad (4.52)$$

Steigt die Strahlenanzahl, müssen demnach auch die Formerhaltungskosten steigen, und zwar überproportional für  $\alpha > 1$ . Für den Fall  $\alpha = 2$  soll der Zusammenhang zwischen der Strahlenanzahl  $n$  und den Formerhaltungskosten  $\lambda$  für das BP-Verfahren auch experimentell untersucht werden.

Dazu wurde der Bereich  $n \in [10, 120]$ ,  $\lambda \in [0.001, 100]$  gleichmäßig abgetastet und für jede Kombination die durchschnittliche Segmentierungsgüte  $\bar{g}$  berechnet. Das Ergebnis ist in Abbildung 4.14 dargestellt. Das globale Maximum befindet sich bei  $n^* = 45$ ,  $\lambda^* = 2.46$ . Dies entspricht einer tatsächlichen Strahlenanzahl von 1311 für die gesamte Oberfläche. Überraschenderweise führen auch deutlich weniger Strahlen zu guten Segmentierungsergebnissen. Bei einer hohen Strahlenanzahl hingegen, ist die Formerhaltung sehr lokal begrenzt, weil nur die Radien benachbarter Strahlen einbezogen werden, um die Komplexität des Verfahrens gering zu halten. Die Folge ist der Verlust an Segmentierungsgenauigkeit.

Zusätzlich wurden für jede Strahlenanzahl die bestmöglichen Formerhaltungskosten  $\lambda^*$  bestimmt. In Abbildung 4.15 sind diese (grüne Punkte) zusammen mit der theoretischen Berechnung durch Gleichung 4.52 ausgehend von  $n = 50$  (blaue Kurve) dargestellt. Auch bei der experimentellen Betrachtung zeigt sich, dass eine höhere Strahlenanzahl eine stärkere Formerhaltung benötigt, und zwar in ähnlicher Weise, wie durch die theoretische Betrachtung erwartet. Die erfolgten Vereinfachungen sind also valide.

Für die drei Segmentierungsverfahren mit unterschiedlicher Optimierung sollen basierend auf den

| Parameter                            |            | Wert(e)             | Bestimmungsweise                      |
|--------------------------------------|------------|---------------------|---------------------------------------|
| <b>Struktur</b>                      |            |                     |                                       |
| Anzahl Strahlen GA                   | $n$        | 293, 413, 213       | Max. der Segmentierungsgüte $\bar{G}$ |
| Anzahl Strahlen BP                   | $n$        | 669, 1259, 1199     | Max. der Segmentierungsgüte $\bar{G}$ |
| Anzahl Strahlen GC                   | $n$        | 2479, 2095, 2321    | Max. der Segmentierungsgüte $\bar{G}$ |
| Maximaler Radius                     | $r_{\max}$ | 30 mm               | Wissen über die max Objektgröße       |
| Knoten pro Strahl                    | $m$        | 120                 | Wissen über Bildauflösung             |
| Saatpunkt                            | $s$        | —                   | Freier Parameter                      |
| <b>Vorverarbeitung</b>               |            |                     |                                       |
| Min Bildintensität                   | $I_{\min}$ | −100 HU             | Wissen über Bildintensitäten          |
| Max Bildintensität                   | $I_{\max}$ | 260 HU              | Wissen über Bildintensitäten          |
| <b>Formerhaltung</b>                 |            |                     |                                       |
| Formerhaltungskosten GA              | $\lambda$  | 0.17, 0.25, 0.25    | Max. der Segmentierungsgüte $\bar{G}$ |
| Exponent GA                          | $\alpha$   | 3.53, 3.19, 4.61    | Max. der Segmentierungsgüte $\bar{G}$ |
| Formerhaltungskosten BP              | $\lambda$  | 5.55, 3.70, 8.33    | Max. der Segmentierungsgüte $\bar{G}$ |
| Exponent BP                          | $\alpha$   | 2.08, 1.46, 2.01    | Max. der Segmentierungsgüte $\bar{G}$ |
| Abweichungstol. GC                   | $\Delta$   | 0.85, 1.15, 1.35 mm | Max. der Segmentierungsgüte $\bar{G}$ |
| <b>Richtungsgew. Log. Grad.</b>      |            |                     |                                       |
| System. Abweichung                   | $\delta$   | 0.38, 0.34, 0.37 mm | Max. der lokalen Erfolgsrate $\nu$    |
| Bildglättung                         | $\sigma$   | 0.12, 0.13, 0.13 mm | Max. der lokalen Erfolgsrate $\nu$    |
| Winkel.Abw.Tol.                      | $\gamma$   | 33.8, 36.1, 32.7 °  | Max. der lokalen Erfolgsrate $\nu$    |
| Maximale Kosten                      | $C_{\max}$ | 5                   | Analyse der Bildgradienten            |
| Grad.Kost. Gewicht GA                | $\rho$     | 0.06, 0.07, 0.08    | Max. der Segmentierungsgüte $\bar{G}$ |
| Grad.Kost. Gewicht BP                | $\rho$     | 5.72, 5.32, 6.29    | Max. der Segmentierungsgüte $\bar{G}$ |
| Grad.Kost. Gewicht GC                | $\rho$     | 3.67, 3.36, 1.92    | Max. der Segmentierungsgüte $\bar{G}$ |
| <b>Kumulierte Intensitäten GA,GC</b> |            |                     |                                       |
| System. Abweichung                   | $\delta$   | 0.81, 0.78, 0.81 mm | Max. der globalen Erfolgsrate $\nu'$  |
| Bildglättung                         | $\sigma$   | 0.24, 0.22, 0.24 mm | Max. der globalen Erfolgsrate $\nu'$  |
| Min Bildintensität                   | $I_{\min}$ | 43 HU               | Grauwert-Histogramm-Analyse           |
| Max Bildintensität                   | $I_{\max}$ | 119 HU              | Grauwert-Histogramm-Analyse           |
| <b>Homogenität BP</b>                |            |                     |                                       |
| System. Abweichung                   | $\delta$   | 0.23, 0.31, 0.21 mm | Max. der globalen Erfolgsrate $\nu'$  |
| Bildglättung                         | $\sigma$   | 2.19, 2.16, 2.20 mm | Max. der globalen Erfolgsrate $\nu'$  |

Tabelle 4.12.: Übersicht aller Segmentierungsparameter und die Art der Bestimmung. Für alle durch Optimierung bestimmten Parameter sind die unterschiedlichen Werte für die Trainingsmengen  $\mathcal{T}_1$ ,  $\mathcal{T}_2$  und  $\mathcal{T}_3$  der Kreuzvalidierung separat angegeben.

vorangegangenen theoretischen und experimentellen Betrachtungen nun die optimalen Parameter  $n^*$  und  $\alpha^*$ ,  $\lambda^*$  bzw.  $\Delta^*$  für die jeweils besten Kostenfunktionen ermittelt werden. Die Anzahl der Strahlen  $n$  bezieht sich von nun an wieder auf die Gesamtanzahl. Die zu maximierende Zielfunktion ist damit für eine Trainingsmenge  $\mathcal{T}$  die durchschnittliche Segmentierungsgüte  $\bar{\mathcal{G}}_{\mathcal{T}}(n, \alpha, \lambda)$  für GA und BP bzw.  $\bar{\mathcal{G}}_{\mathcal{T}}(n, \Delta)$  für GC. Die Optimierung findet, wie in den Parameteroptimierungsschritten zuvor, durch das Simplex-Optimierungsverfahren statt, wobei zuvor eine Parametersubstitution analog zur Gleichung 4.38 für alle Parameter zum Einsatz kommt. Wie Abbildung 4.14 suggeriert, ist solch ein lokales Optimierungsverfahren aufgrund der Konvexität der Funktion geeignet, das globale Optimum zu finden. Um eine aussagekräftige Evaluierung zu ermöglichen, wurden die Parameter im Zuge der 3-fachen Kreuzvalidierung, wie zuvor auch schon, unabhängig auf den drei Trainingsmengen bestimmt.

Die erhaltenen Parameter sind zusammen mit allen anderen Parametern in der Übersichtstabelle 4.12 eingetragen. Auch in diesem letzten Optimierungsschritt wurden konsistente Werte erzielt. Die Anzahl der Strahlen variiert für die unterschiedlichen Verfahrensarten beträchtlich. Es hat sich herausgestellt, dass GA deutlich weniger Strahlen als BP und BP wiederum deutlich weniger Strahlen als GC für die jeweils bestmögliche Segmentierungsgüte benötigt.

Damit sind alle Parameter der Segmentierungsverfahren bestimmt. Die in Tabelle 4.12 eingetragenen Parameter stellen die Grundlage für die ausführliche Evaluation in Abschnitt 5.1.4 dar.

### 4.1.8. Erweiterungen

Das in den vorangegangenen Abschnitten beschriebene Segmentierungsverfahren erfüllt bereits einen Großteil der in Abschnitt 1.3 formulierten Anforderungen. Folgende relevante Problemstellungen blieben allerdings bisher unberücksichtigt:

- Das Segmentierungsergebnis ist abhängig von der Wahl des Saatpunkts und damit auch vom Anwender. Zur Erhöhung der Robustheit sollte es aber so weit wie möglich unabhängig von der manuellen Initialisierung sein.
- Die durch Gleichung 4.17 beschriebene Formerhaltungsenergie ist nur für Kugeln minimal. Die Form von Lymphknoten kann davon aber beträchtlich abweichen. Eine geeignetere Formerhaltung ist wünschenswert.
- Durch mangelnden Bildkontrast, Rauschen oder andere unvorhergesehene Faktoren, kann es passieren, dass das Segmentierungsergebnis immer noch nicht zufriedenstellend ist. Eine komfortable schnelle Art der interaktiven Korrektur wäre für solche Fälle wünschenswert.

Im Folgenden werden drei Erweiterungen des Radialstrahl-basierten Lymphknotensegmentierungsverfahrens vorgestellt.

**Invarianz gegenüber des Saatpunkts:** Der Saatpunkt sollte innerhalb des Lymphknotens liegen, so dass die gesamte Lymphknotengrenze auf direktem Weg vom Saatpunkt erreichbar ist, ohne dass sie an anderer Stelle geschnitten wird. Der Bereich, auf den dies zutrifft, wird als das Zentrum der sternkonvexen Struktur bezeichnet. In vielen Fällen, jedoch nicht notwendigerweise, liegt der Schwerpunkt des Objektes innerhalb des Zentrums. Aus diesem Grund und weil vom Schwerpunkt des Objektes ausgehend die Objektoberfläche in alle Richtungen so gleich weit wie möglich entfernt ist,



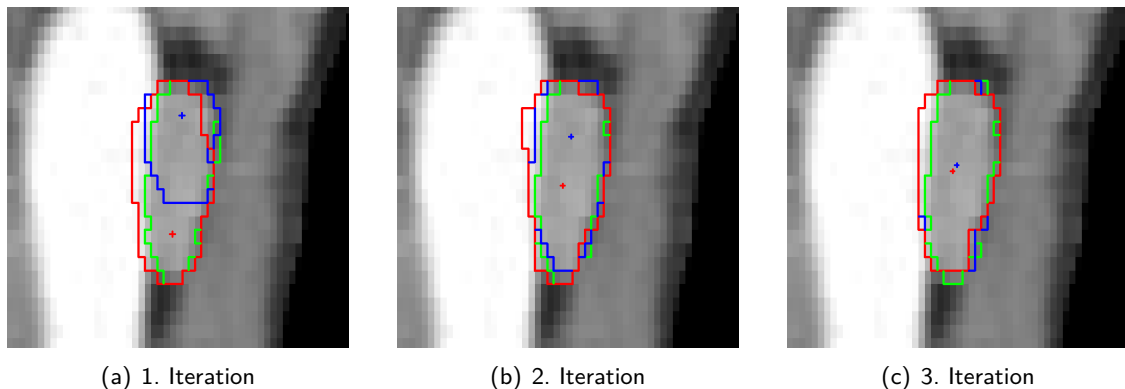
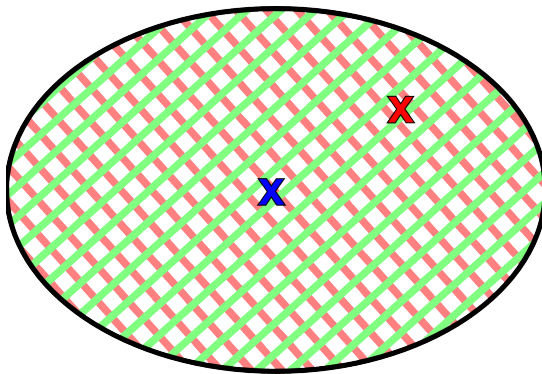


Abbildung 4.16.: Trotz der Wahl unterschiedlicher Saatpunkte (rot und blau) wird durch die Neuberechnung des Saatpunkts nach einigen Iterationen ein sehr ähnliches Segmentierungsergebnis erzielt.

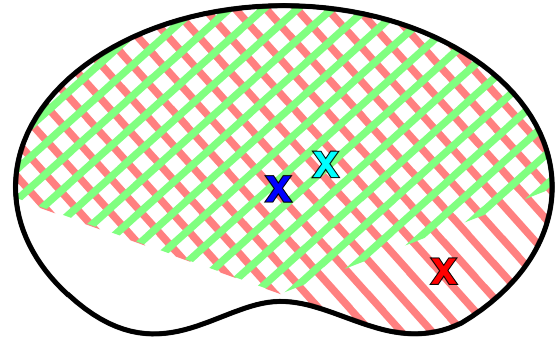
eignet sich der Schwerpunkt besonders als Saatpunkt. Eine manuelle Initialisierung des Saatpunkts wird in den seltensten Fällen mit dem Schwerpunkt übereinstimmen, was eine gewisse Intra- und Inter-Observer-Variabilität nach sich zieht.

Unter der Annahme, dass die Segmentierungsergebnisse bei unterschiedlichen Saatpunkten in der Nähe des Zentrums ähnlich sind (siehe Abschnitt 5.1.4 für eine experimentelle Bestätigung dieser Annahme), kann ein iteratives Verfahren eine gewisse Invarianz gegenüber des Saatpunkts erreichen: Zunächst wird das Segmentierungsverfahren wie gewohnt an einem manuell initialisierten Saatpunkt  $s_0$  angewendet und eine erste Segmentierung  $S_0$  wird erhalten. Daraus wird dann der Masseschwerpunkt bestimmt und als initialer Saatpunkt  $s_1$  für eine wiederholte Anwendung des Segmentierungsverfahrens verwendet.  $s_1$  sollte nun näher an dem tatsächlichen Schwerpunkt des Lymphknotens liegen. Dieser Vorgang wird nun wiederholt, bis eine maximale Anzahl an Iterationen erreicht wurde oder der Schwerpunkt konvergiert ist (siehe Abbildung 4.16).

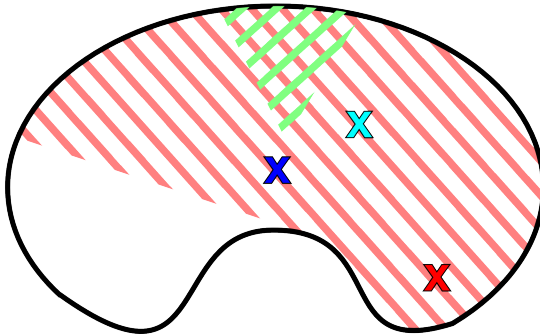
Diese Heuristik hat Potential zur Erhöhung der Robustheit. Konvergenz ist allerdings nicht für alle unterstützten Formen (vgl. Abschnitt 4.1.3) und auch nicht für den Fall von Ungenauigkeiten der Segmentierung garantiert. Im Falle konvexer Strukturen (siehe Abbildung 4.17(a)) kann jedoch für einen perfekten Radialstrahl-basierten Segmentierungsansatz Konvergenz einfach gezeigt werden: Da in einer konvexen Menge jeder Punkt mit jedem anderen Punkt verbunden werden kann, so dass alle Punkte der Verbindungsstrecke sich ebenfalls innerhalb der Menge befinden, führt eine Segmentierung von einem beliebigen Punkt innerhalb der Menge ausgehend zu einer Segmentierung der gesamten Menge. Der Schwerpunkt dieser perfekten Segmentierung fällt natürlich mit dem Objektschwerpunkt zusammen und das Verfahren konvergiert. Notwendige Bedingung für die Konvergenz aller unterstützten Strukturen (sternkonvexe Strukturen mit glatter Oberfläche) ist, dass der Schwerpunkt der Struktur sich innerhalb des Zentrums – also dem Bereich, von dem aus alle Punkte der gesamten Struktur *erreicht* werden können – befindet. Dann kann das gesamte Objekt segmentiert werden und der Schwerpunkt fällt wieder auf dieselbe Stelle. Hinreichende Bedingung ist jedoch, dass sich der initiale Saatpunkt ebenfalls innerhalb des Zentrums befindet, oder zumindest nach endlich vielen Iteration des Verfahrens sich in das Zentrum hinein bewegt (siehe Abbildung 4.17(b)). Befindet sich der Schwerpunkt der Struktur außerhalb des Zentrums, aber innerhalb des



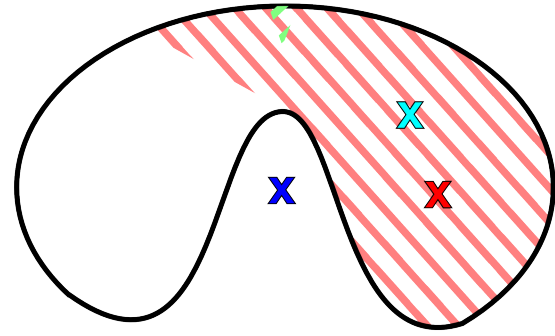
(a) Bei konvexen Strukturen führt ein beliebiger Saatpunkt innerhalb der Struktur zu einer Segmentierung der gesamten Struktur. Der Schwerpunkt der Segmentierung fällt somit stets mit dem Schwerpunkt der Struktur zusammen und das Verfahren konvergiert.



(b) Auch bei annähernd konvexen Strukturen fällt der Schwerpunkt der Segmentierung stets ins Zentrum. Da sich hier der Strukturschwerpunkt innerhalb des Zentrums befindet, konvergiert das Verfahren auch in diesem Fall für alle Saatpunkte.



(c) Liegt der Schwerpunkt der Struktur nicht im Zentrum, kann Konvergenz weder ausgeschlossen noch garantiert werden. In jedem Fall ist mit Segmentierungsungenauigkeiten zu rechnen.



(d) Liegt der Schwerpunkt der Struktur außerhalb der Struktur, ist das Saatpunkt-invariante Verfahren ungeeignet.

Abbildung 4.17.: Unterschiedliches Konvergenzverhalten des Saatpunkt-invarianten Verfahrens für unterschiedliche sternkonvexe Formen unter der Annahme eines perfekten Radialstrahl-basierten Segmentierungsverfahrens. Es werden jeweils Schwerpunkt der Struktur (blau), das Zentrum der Struktur (grüne Schraffierung), ein Saatpunkt  $s_0$  (rot), die dazugehörige bestenfalls erzielbare Segmentierung  $S_0$  (rote Schraffierung) und der Schwerpunkt dieser Segmentierung  $s_1$  (türkis) dargestellt.

Objektes (siehe Abbildung 4.17(c)), ist Konvergenz nicht garantiert. Sobald der Saatpunkt das Zentrum erreicht, wird die gesamte Struktur segmentiert, was zur Folge hat, dass der Saatpunkt sich wieder aus dem Zentrum hinaus begibt und in der nächsten Iteration es prinzipbedingt zu Segmentierungsungenauigkeiten kommt. Wohin der Schwerpunkt dann fällt kommt auf die konkrete Form an. Konvergenz kann weder ausgeschlossen, noch garantiert werden. Für den Fall, dass der Schwerpunkt sich sogar außerhalb der Struktur befindet (siehe Abbildung 4.17(d)), ist das Verfahren gänzlich ungeeignet, weil ein Saatpunkt außerhalb des Objektes ein unklares Segmentierungsergebnis nach

sich zieht. Solche Strukturen sind aber aufgrund der Glattheitsanforderung unwahrscheinlich. Neben diesen theoretischen Überlegungen, die von einem perfekten Radialstrahl-basierten Segmentierungsverfahren ausgehen, ist eine experimentelle Validierung des Verfahrens in der praktischen Anwendung nötig. Diese ist in Abschnitt 5.1.4 zu finden.

**Bildgradientenbasiertes Formwissen:** In diesem Abschnitt wird das in [11] veröffentlichte Formwissen vorgestellt. Das bisher verwendete Formwissen fordert, dass benachbarte Strahlen einen möglichst ähnlichen Radius haben. Die zugrundeliegende Annahme ist dabei, dass die Form der einer Kugel ähnlich ist und dass der Saatpunkt in der Nähe des Zentrums liegt. Je weniger diese Annahme zutrifft, desto geringer darf der Einfluss des Formwissens beim Segmentierungsprozess sein, um dennoch eine Segmentierung der Zielstruktur erreichen zu können. Im Gegenzug vergrößert eine Abschwächung des Formwissens allerdings die Gefahr des Auslaufens der Segmentierung in benachbarte Strukturen. Wie in Abschnitt 2.1.1 beschrieben, trifft die Annahme einer Kugel für Lymphknoten nicht sonderlich gut zu. Daher wird ein Formwissen benötigt, welches eine größere Formvariabilität unterstützt.

Dazu wird zunächst der Verlauf von Objektgrenzen untersucht. Bei der Betrachtung der Bildgradientenrichtung an der Objektgrenze fällt auf, dass diese in die Normalenrichtung der Objektoberfläche zeigt. In einer kleinen Umgebung um einen Punkt auf der Oberfläche kann diese als Ebene approximiert werden, deren Normalenrichtung der Bildgradientenrichtung  $\nabla I(\mathbf{x})$  an dieser Stelle  $\mathbf{x}$  entspricht.

Nun soll ein Radialstrahl-basiertes Segmentierungsverfahren betrachtet werden. Sind genügend Strahlen vorhanden, kann angenommen werden, dass die Objektoberflächenpunkte benachbarter Strahlen sehr nahe aneinander liegen und die gerade beschriebene Approximation der Objektoberfläche durch eine Ebene zutreffend ist. Hat nun der Strahl  $i$  den Oberflächenpunkt  $\mathbf{x}_i$ , sollte der Oberflächenpunkt auf dem benachbarten Strahl  $j$  sich möglichst nahe am Schnittpunkt  $\mathbf{y}_{ij}$  des Strahls  $j$  mit der Ebene durch  $\mathbf{x}_i$  und der Normalenrichtung  $\nabla I(\mathbf{x}_i)$  befinden (siehe Abbildung 4.18(a)). Durch Einsetzen der Gleichung des Strahls in die Hesseschen Normalform der Ebenengleichung kann dieser Schnittpunkt berechnet werden:

$$\mathbf{y}_{ij} = \mathbf{s} + \frac{\nabla(G_\sigma \star I(\mathbf{x}_i)) \cdot \mathbf{d}_i}{\nabla(G_\sigma \star I(\mathbf{x}_i)) \cdot \mathbf{d}_j} r_i \mathbf{d}_j. \quad (4.53)$$

Der gewünschte Radius  $\tau_j$  des Strahls  $j$  kann dann als Funktion des Radius von Strahl  $i$  aufgefasst werden:

$$\tau_j(r_i) = \begin{cases} \frac{\nabla(G_\sigma \star I(\mathbf{s} + r_i \mathbf{d}_i)) \cdot \mathbf{d}_i}{\nabla(G_\sigma \star I(\mathbf{s} + r_i \mathbf{d}_i)) \cdot \mathbf{d}_j} r_i & ; |\nabla(G_\sigma \star I(\mathbf{s} + r_i \mathbf{d}_i))| > g_{\min} \\ r_i & ; \text{sonst} \end{cases} \quad (4.54)$$

wobei bei einem unscharfen Übergang oder geringem Bildkontrast, also kleinen Bildgradientenstärken, darauf verzichtet und angenommen wird, dass der gewünschte Radius dem des Nachbarstrahls entspricht. Der Grund dafür ist, dass schwache Bildgradienten nicht notwendigerweise von einer Objektgrenze, sondern möglicherweise von Bildstörungen stammen und deren Richtungen damit nicht aussagekräftig sind. Der Parameter  $g_{\min}$  gibt die minimale Bildgradientenstärke an, die notwendig ist, damit die Richtung als aussagekräftig aufgefasst wird.

Die Gleichung der Formerhaltung 4.17 kann nun entsprechend erweitert werden:

$$E_{\text{Form}}(\mathbf{l}) = \sum_{\forall j \in \mathcal{N}(i)} \left| \left( \tau_i \left( \frac{r_{\max}}{m} l_i \right) - \frac{r_{\max}}{m} l_j \right) \right|^\alpha. \quad (4.55)$$

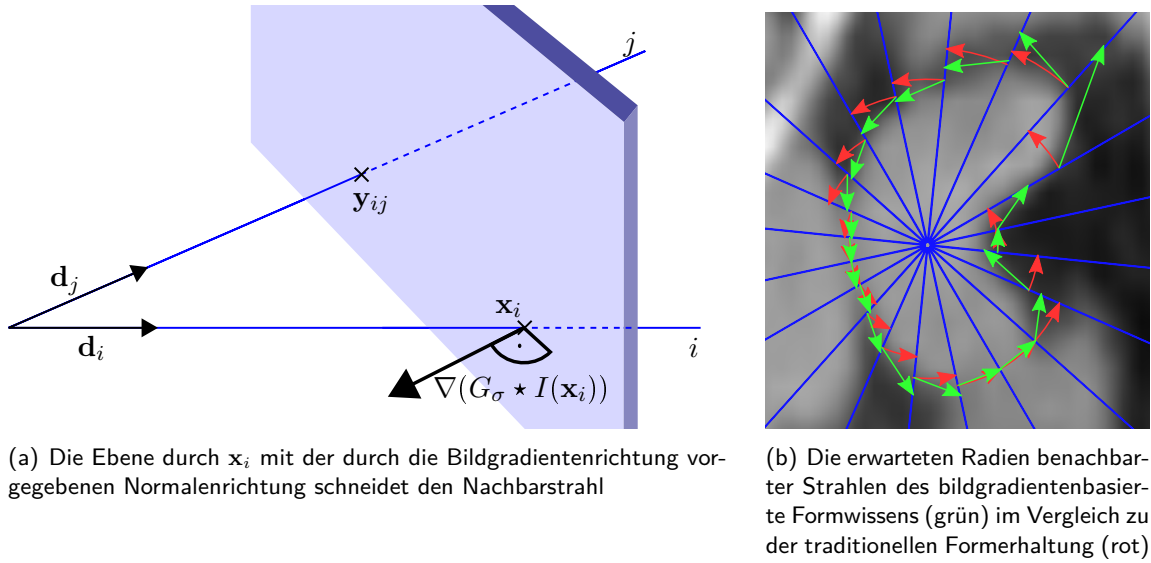


Abbildung 4.18.: Der erwartete Radius des benachbarten Strahls ergibt sich aus der Bildgradientenrichtung

Das Einbeziehen der Bildgradientenrichtung führt dazu, dass diese Art der Formerhaltung nicht kugelförmige Strukturen, sondern Strukturen, welche tatsächlich im Bild vorliegen, bevorzugt. Ein Beispiel dazu ist in Abbildung 4.18(b) für einen typischen bohnenförmigen Lymphknoten zu sehen.

Die Kehrseite dieser Art der Formerhaltung besteht allerdings darin, dass die durch Gleichung 4.17 gegebene Konvexität der ursprünglichen Formerhaltung nicht mehr gegeben ist und damit die Konvergenz des Belief-Propagation-Algorithmus nicht mehr garantiert ist. Daher muss die Sinnhaftigkeit experimentell nachgewiesen werden, was in Abschnitt 5.1.4 geschieht.

Die Parameter  $\sigma$  und  $g_{\min}$  wurden durch informelle Experimente auf  $0.5 \text{ mm}$  bzw.  $80$  festgelegt. Die Änderung der Formerhaltung hat auch einen Einfluss auf andere Parameter, insbesondere auf den Exponenten  $\alpha$  und den Formerhaltungskostenparameter  $\lambda$ . Daher werden diese beiden Parameter, wie schon in den vergangenen Abschnitten beschrieben, durch Maximierung der durchschnittlichen Segmentierungsgüte  $\bar{G}(\alpha, \lambda)$  für jede der drei Trainingsmengen gesondert bestimmt.  $\alpha^*$  ergibt sich zu  $2.25, 2.13, 1.52$  und  $\lambda^*$  zu  $5.55, 3.70, 5.00$ . Im Vergleich zum Ausgangsverfahren (BP mit Formerhaltung nach Gleichung 4.17) sind durch das neue Formerhaltungswissen die Formerhaltungskosten  $\lambda$  oder der Formerhaltungsexponent  $\alpha$  größer. Die hier vorgestellte, auf dem Bildgradienten basierende Formerhaltung hat also einen größeren Einfluss auf die zu minimierende Energie als beim Standardformwissen. Diese Beobachtung spricht dafür, dass ein durch die Bildgradientenrichtung bestimmter erwarteter Radius eines Nachbarstrahls im Fall von Lymphknoten tatsächlich zutreffender ist, als die Erwartung, dass benachbarte Strahlen einen möglichst ähnlichen Radius haben.

**Interaktive Segmentierung:** Um eine manuelle Korrektur durch einen Experten während des Segmentierungsprozesses zu ermöglichen, wird in diesem Abschnitt das in [12] veröffentlichte Verfahren zur interaktiven Segmentierung vorgestellt. Grundlage für das interaktive Verfahren ist das in Abschnitt 4.1.4 beschriebene auf Belief-Propagation basierende Verfahren mit der im vergangenen

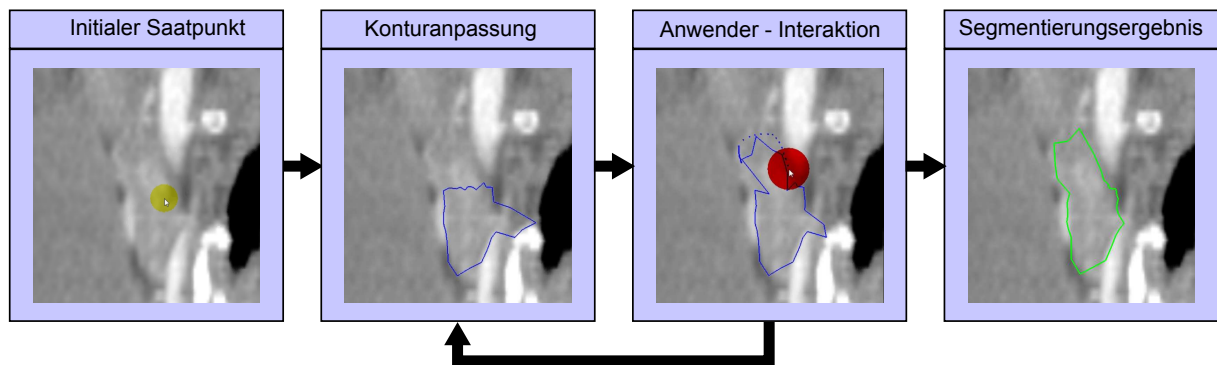


Abbildung 4.19.: Ablauf der interaktiven Segmentierung eines kontrastarmen Lymphknotens aus Sicht des Anwenders

Abschnitt beschriebenen Formerhaltung. Obwohl GC etwas bessere Werte als BP erzielt hat, kann dieses Verfahren nicht verwendet werden, weil die Berechnung der optimalen Belegung nicht iterativ erfolgt. Der iterative Charakter von BP eignet sich jedoch zum interaktiven Eingriff des Anwenders.

Aus Sicht des Anwenders läuft die Segmentierung eines Lymphknotens unter Verwendung einer graphischen Benutzeroberfläche in folgenden Schritten ab:

1. Der Anwender wählt einen Saatpunkt durch einen Mausklick in das ungefähre Zentrum des zu segmentierenden Lymphknotens.
2. Das System zeigt ein 3D-Modell der aktuell am wahrscheinlichsten angenommenen Kontur in einer vom Anwender ausgewählten axialen, coronalen und sagittalen Schicht an. Ein im Hintergrund ablaufender Prozess aktualisiert kontinuierlich diese Kontur.
3. Der Anwender inspiziert die angezeigte Kontur in beliebigen Schichten.
4. Besteht Korrekturbedarf, zeichnet der Anwender die gewünschte Lymphknotenkontur in einer beliebigen Schicht ein. Das System bezieht diese eingezeichneten Punkte in den Konturberechnungsprozess ein.
5. Ist das gewünschte Segmentierungsergebnis erzielt, bestätigt der Anwender dieses.

Dieses Vorgehen ist in Abbildung 4.19 noch einmal verdeutlicht. Da der Algorithmus einem Radialstrahl-basierten Ansatz zugrunde liegt, ist die Klasse der unterstützten Strukturen weiterhin auf sternkonvexe Formen beschränkt.

Um das beschriebene Verhalten zu erzielen, wird das BP-Verfahren folgendermaßen erweitert:

- Nach jedem Iterationsschritt  $t$ , in dem die Strahlen mit ihren Nachbarstrahlen Nachrichten austauschen, wird die gerade wahrscheinlichste Belegung  $I_t^*$  berechnet, ein korrespondierendes Polygonnetz erstellt und visualisiert. Bei dem ursprünglichen BP-Verfahren findet dieser Vorgang nur einmal nach der Konvergenz des Verfahrens oder bei Erreichen einer maximalen Iterationsanzahl statt.
- Die Terminierungsbedingungen des Algorithmus werden durch eine manuelle Bestätigung ersetzt. Ohne Eingriff seitens des Anwenders terminiert das Schicken der Nachrichten zwischen benachbarten Strahlen also nicht.

- Beim Einzeichnen eines gewünschten Konturpunktes  $\mathbf{x}^*$  durch den Anwender wird die Kostenfunktion  $C$  dahingehend verändert, dass dem Punkt  $\mathbf{x}^*$  minimale Kosten zugeordnet werden. Dazu wird zunächst derjenige Strahl  $k$  ausgewählt, dessen Richtung  $\mathbf{d}_k$  am nächsten zu  $\mathbf{x}^* - \mathbf{s}$  ist:

$$k = \underset{\forall i}{\operatorname{argmin}} \angle(\mathbf{d}_i, \mathbf{x}^* - \mathbf{s}) . \quad (4.56)$$

Die entsprechende gewünschte Belegung  $l_k^*$  ergibt sich zu:

$$l_k^* = \lfloor \frac{|\mathbf{x}^* - \mathbf{s}|}{r_{\max}} m \rfloor . \quad (4.57)$$

Die neuen Kosten des Strahls  $k$  ergeben sich dann zu

$$C_k(j) = \begin{cases} 0 & ; j = l_k^* \\ C_{\max} & ; \text{sonst} \end{cases} . \quad (4.58)$$

In den folgenden Iterationsschritten werden die Nachrichten dann dementsprechend automatisch angepasst, so dass weiterhin eine glatte Kontur entsteht.

##### 4.1.9. Zusammenfassung

**Methoden:** Nach einigen Experimenten mit einem größeninvarianten Masse-Feder-Modell und einer vielversprechenden Radialstrahl-basierten 2D-Segmentierungsmethode wurde eine Radialstrahl-basierte 3D-Segmentierungsmethode entwickelt, welche zur Segmentierung einzelner Lymphknoten aus klinischen CT-Daten eingesetzt werden kann.

Ausgehend von einem Saatpunkt werden Strahlen radial in alle möglichen Richtungen gleichverteilt geschickt. Eine Segmentierung wird durch eine geschlossene Oberfläche um den Saatpunkt beschrieben, welche durch je einen Radius für jeden Strahl parametrisiert ist. Diese Radien werden von einem Optimierungsverfahren unter Einbeziehung von lokalem Formwissen und Wissen über das Aussehen von Lymphknoten in CT-Aufnahmen in Form von Kostenfunktionen gefunden.

**Eigene Beiträge:** Im Vergleich zu existierenden Lymphknotensegmentierungsverfahren ergeben sich folgende Neuerungen:

- **Vergleich von Optimierungsverfahren:** In existierenden Radialstrahl-basierten Verfahren wird entweder das lokale Gradientenabstiegsverfahren oder das globale Graph-Cut-Verfahren zum Finden des Radienvektors eingesetzt. Im Rahmen dieser Arbeit wurde zum ersten Mal Belief-Propagation verwendet, welches durch den iterativen Charakter eine interaktive Erweiterung und eine flexible Art der Einbeziehung von Formwissen ermöglicht. Während in der Literatur jeweils ein bestimmtes Optimierungsverfahren verwendet wird, findet hier zum ersten Mal ein Vergleich von Optimierungsverfahren statt, wobei das Graph-Cut Verfahren zu den besten Segmentierungsergebnissen führt.
- **Vergleich von Kostenfunktionen:** Anstatt lediglich eine einzige Kostenfunktion zu motivieren und vorzustellen, werden unterschiedliche bildgradienten- und intensitätsbasierte Kostenfunktionen vorgestellt und gegenübergestellt. Für Lymphknoten findet solch ein Vergleich von Kostenfunktionen zum ersten Mal statt.

- **Neuartige Kostenfunktion.** Die in dieser Arbeit verwendeten bildbasierten Kostenfunktionen sind größtenteils Erweiterungen existierender Kostenfunktionen. Die auf kumulierten Intensitäten beruhende Kostenfunktion ist allerdings eine Neuentwicklung. In Kombination mit einer bildgradientenbasierten Kostenfunktion erweist diese sich im Vergleich zu existierenden intensitätsbasierten Kostenfunktionen als besser geeignet.
- **Bestimmung der Parameter:** Anstatt die Parameter heuristisch festzulegen, wurde in dieser Arbeit ein Verfahren entwickelt, welches auf Grundlage von Referenzsegmentierungen optimale Parameter automatisch bestimmt.
- **Saatpunktinvariantes Verfahren:** Zur Erhöhung der Robustheit des Segmentierungsergebnisses gegenüber dem manuell gesetzten Saatpunkt wurde im Rahmen dieser Arbeit ein Verfahren angewendet, welches den Saatpunkt iterativ anpasst.
- **Bildgradientenbasiertes Formwissen:** Ausnahmslos fordern existierende Radialstrahl-basierte Verfahren zur Formerhaltung, dass benachbarte Strahlen möglichst ähnliche Radien aufweisen, was zu einer Bevorzugung von kugelförmigen Strukturen führt. In dieser Arbeit wurde zum ersten Mal ein alternatives lokales Formwissen vorgestellt, welches mit Belief-Propagation als Optimierungsverfahren im Radialstrahl-basierten Lymphknotensegmentierungsverfahren zum Einsatz kommt. Anstatt Kugeln bevorzugt es tatsächlich im Bild vorhandene Strukturen.
- **Interaktives Verfahren:** Zum ersten Mal wurde ein Radialstrahl-basiertes Segmentierungsverfahren um einen interaktiven Modus erweitert, welcher es einem Anwender während des Optimierungsprozesses durch Belief-Propagation erlaubt, das Segmentierungsergebnis zu beeinflussen.

**Vorteile:** Das entwickelte Radialstrahl-basierte Verfahren ist in der Lage, die großen Herausforderungen der Lymphknotensegmentierung zu meistern. Durch den Einsatz eines globalen Optimierungsverfahrens werden alle Lymphknotengrößen gleichermaßen unterstützt und abgesehen von dem Saatpunkt ist keine weitere Initialisierung nötig. Existierende Verfahren benötigen oftmals mehr Anwenderinteraktion zur Initialisierung. Durch die Verwendung von Formwissen können gegenüber rein bildbasierten Segmentierungsverfahren auch Lymphknoten mit geringerem Bildkontrast an der Objektgrenze erfolgreich segmentiert werden. Da es sich dabei um lokales anstatt globales Formwissen handelt, kann eine hohe Formvarianz, wie sie typisch bei Lymphknoten ist, abgedeckt werden, ohne dabei eine bestimmte globale Form a priori zu bevorzugen. Darüber hinaus wurde bei der Entwicklung der Kostenfunktionen darauf geachtet, dass auch pathologische Lymphknoten mit zentralem nekrotischen Gewebe unterstützt werden.

**Nachteile und Lösungsansätze:** Der Nachteil aller Radialstrahl-basierten Verfahren ist die prinzipbedingte Beschränkung auf sternkonvexe Formen und die ungleichmäßige Auflösung des Segmentierungsergebnisses. Bereiche, die näher an dem Saatpunkt liegen, haben eine deutlich höhere Auflösung als weiter außen liegende Bereiche. Eine Lösungsansatz dafür ist die Verwendung eines modellbasierten Segmentierungsverfahrens, welches durch das radialstrahlbasierte Segmentierungsverfahren initialisiert wird. So kommen die Nachteile beider Verfahren nicht zum Tragen: das kombinierte Verfahren hat durch den modellbasierten Anteil eine gleichmäßige Auflösung und ist nicht auf sternkonvexe Strukturen beschränkt. Gleichzeitig hat es durch den radialstrahlbasierten Anteil einen großen Konvergenzbereich.

Ein weiterer Nachteil ist die Abhängigkeit von dem manuell platzierten Saatpunkt. Dadurch kann dieses Verfahren zum einen nicht vollautomatisch im Hintergrund ablaufen, sondern erfordert Anwenderinteraktion, wenn auch nur in geringem Maße. Zum anderen führt dieser Einfluss zu einer Abhängigkeit vom Anwender. In Abschnitt 5.1.4 wird den praktischen Auswirkungen dieser Nachteile experimentell nachgegangen. Eine Möglichkeit, die Anwenderabhängigkeit zu vermeiden, besteht darin, geeignete Saatpunkte durch ein automatisches Lymphknotendetektionsverfahren, wie sie z.B. in [64][65][66] oder [67] vorgestellt werden, zu bestimmen.



## 4.2. Registrierung von Lymphknoten in Kopf-Hals-Tomographien

Eine präzise Registrierung von Lymphknoten in CT- und MRT-Aufnahmen ist unter anderem zur Bildfusion für eine anschließende Bestimmung von Merkmalen nötig. Eine rigide Registrierung der gesamten Aufnahmen ist allerdings nicht ausreichend, weil schon eine leicht unterschiedliche Position des Patienten in beiden Aufnahmen zu einem lokalen Unterschied führt, der nicht durch eine globale Transformation ausgeglichen werden kann. Übliche deformierbare Registrierungsverfahren, wie z.B. die B-Spline-Registrierung, sind ebenfalls nicht zielführend, weil diese die Lymphknoten potentiell deformieren. Durch die komplementäre Information von multimodalen Aufnahmen könnte es bspw. passieren, dass ein ganzer Lymphknoten in der MRT-Aufnahme lediglich auf das innere nekrotische Gewebe in der CT-Aufnahme abgebildet wird. Darüber hinaus ist bei der Anwendung auf Zeitserien eine vergleichende Beurteilung desselben Lymphknotens über mehrere Aufnahmen hinweg nicht möglich, weil Unterschiede in der Anatomie von der Registrierung als Deformation aufgefasst und ausgeglichen werden. Aus diesen Gründen wird in dieser Arbeit ein lokal rigides Registrierungsverfahren vorgestellt, welches lediglich den Lymphknoten und etwas umliegendes Gewebe betrachtet. Die Lokalisation des Lymphknotens ist durch die Segmentierung mit dem im vorangegangenen Kapitel vorgestellten Verfahren aus der CT-Aufnahme gegeben. Da der Konvergenzbereich eines solchen Verfahrens durch die prinzipbedingte Lokalität sich allerdings nur auf einen kleinen Bereich beschränkt, ist davor eine ganzheitliche Registrierung der Aufnahmen nötig, dessen Genauigkeit nicht allzu groß sein muss, dafür aber einen großen Konvergenzbereich haben muss und robust gegenüber Bildartefakten sein soll. Dazu wird ein Regmentierungsverfahren (siehe Kapitel 2.2.3) vorgestellt, welches aus 5 Teilschritten besteht und auf einem artikulierten Atlas des Skeletts basiert:

- In einem ersten Schritt wird der Kopf rigide registriert.
- Anschließend erfolgt eine Anpassung des artikulierten Atlas an die CT-Aufnahme, wodurch i) eine Segmentierung der einzelnen rigiden Strukturen (Knochen) erfolgt und ii) der artikulierte Atlas durch personalisiertes Wissen angereichert wird.
- Letzteres ermöglicht es, den Atlas in einem weiteren Schritt an die andere Bildmodalität anzupassen, um eine rigide Transformation für jeden Knochen zu erhalten.
- Auf Grundlage der nun bekannten Transformationen der einzelnen Knochen wird daraufhin ein dichtes Deformationsfeld berechnet. Optional kann auch die Hautoberfläche als weitere korrespondierende Landmarke in die Berechnung des Deformationsfeldes einbezogen werden.
- Das entstandene Deformationsfeld wird nun zur Initialisierung des lokal rigiden Lymphknoten-registrierungsverfahrens verwendet.

Der gesamte Ablauf der Registrierung von Lymphknoten ist in Abbildung 4.20 dargestellt und die einzelnen Schritte werden in den nächsten Abschnitten ausführlich beschrieben. Das Gesamtkonzept der Methoden zur globalen Registrierung des Kopf-Hals-Bereichs wurde in [13] veröffentlicht.

### 4.2.1. Multimodale rigide Registrierung des Kopfes

Zur bildbasierten multimodalen Registrierung des Kopfes hat sich die Maximierung der *Mutual Information* [72][73] beider Aufnahmen durchgesetzt [98][99]. Die Mutual Information  $I(A, B)$  zweier

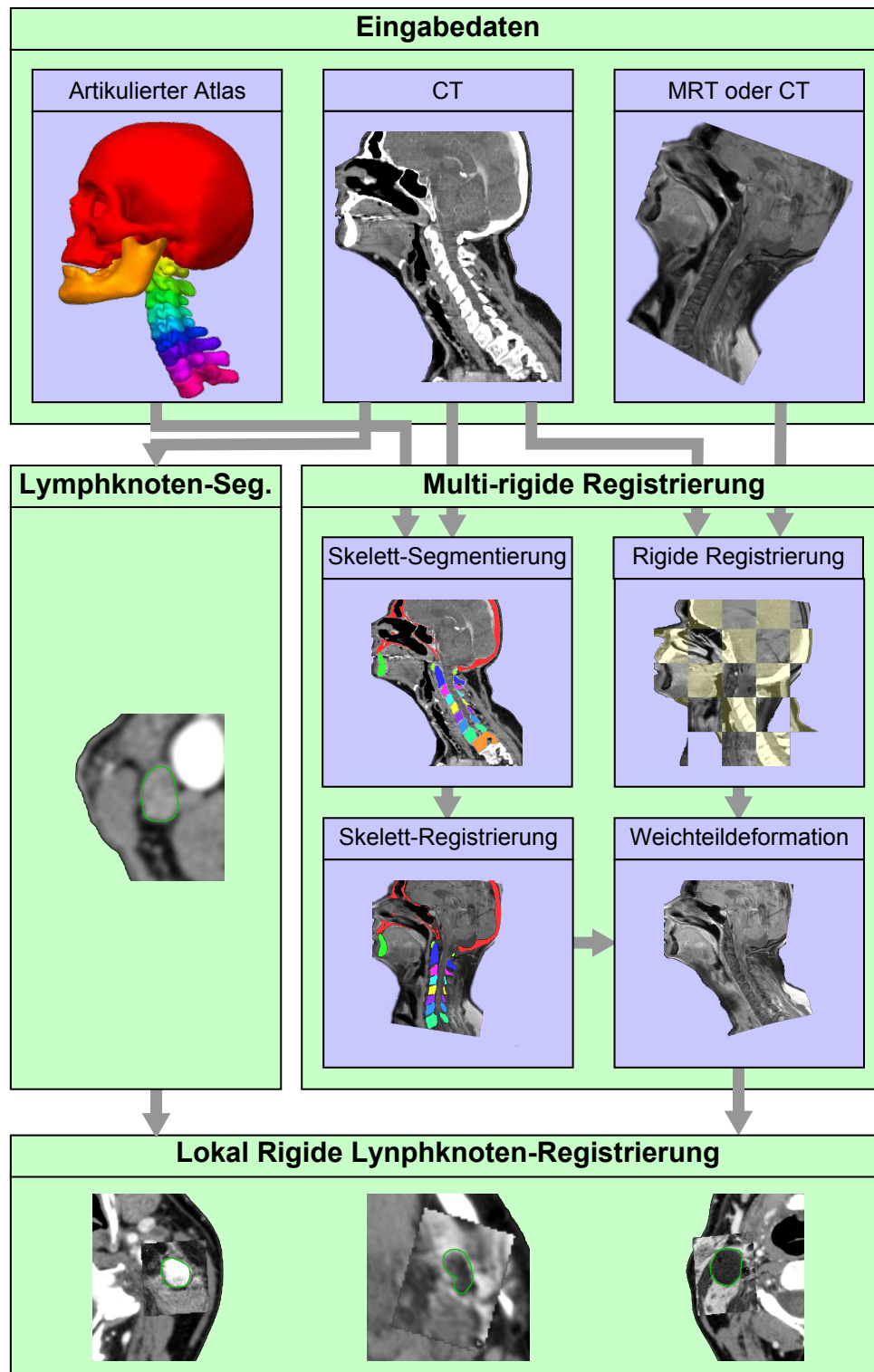


Abbildung 4.20.: Zusammenhang der einzelnen Schritte zur Registrierung von Lymphknoten

Zufallsgrößen (Bilder)  $A$  und  $B$  berechnet sich durch

$$I(A, B) = H(A) + H(B) - H(A, B) , \quad (4.59)$$

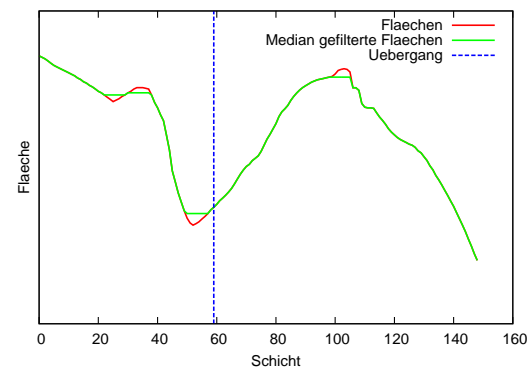
wobei  $H(\cdot)$  die Entropie, also der Informationsgehalt einer Zufallsgröße und  $H(\cdot, \cdot)$  die gemeinsame Entropie zweier Zufallsgrößen ist, welche aus dem überlappenden Bereich der Bilder berechnet wird. Bei der Anwendung in der Bildregistrierung wird die Annahme getroffen, dass die statistische Abhängigkeit der Intensitätswerte beider Bilder bei korrekter Ausrichtung der Aufnahmen am größten ist und damit deren gemeinsame Entropie am geringsten ist. Die Minimierung der gemeinsamen Entropie für sich ist allerdings nicht zielführend, weil Minima vor allem dann auftreten, wenn sich überwiegend homogene Bereiche überlappen (bspw. Hintergrund auf Hintergrund). Da in diesem Fall die Einzelentropien sehr gering sind, verhindert deren Einfluss in die Mutual Information solch ein Minimum.

Bei der Anwendung dieses Prinzips auf klinische Kopf-Hals-Tomographien ergeben sich allerdings folgende Fallstricke:

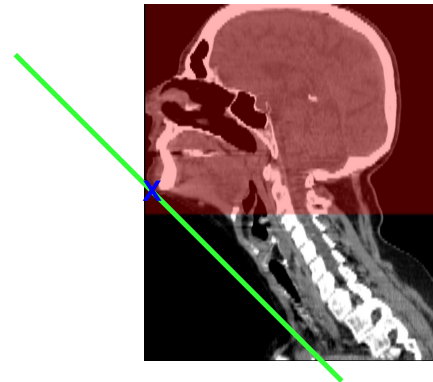
- Die initiale Position und Orientierung der Bilder muss sich bereits in der Nähe des korrekten Registrierungsergebnisses befinden. Andernfalls sind Optimierungsverfahren nicht in der Lage, das korrekte lokale Maximum des Bildähnlichkeitsmaßes zu finden. Eine gängige Methode zur Initialisierung ist die Ausrichtung beider Bilder an deren Schwerpunkten und Hauptachsen [144]. Für MRT-Aufnahmen kann dies im Allgemeinen jedoch nicht angewendet werden, weil diese in der klinischen Praxis oftmals nur einen kleinen Ausschnitt des Kopf-Hals-Bereichs abbilden. Der Schwerpunkt solch einer MRT-Aufnahme fällt daher also im Allgemeinen nicht mit dem einer CT-Aufnahme des gesamten Kopf-Hals-Bereichs zusammen und auch die Ausrichtung an den Hauptachsen ist nicht verlässlich. Es wird also eine andere Initialisierung benötigt, die mit abgeschnittenen Aufnahmen zurecht kommt.
- Im Gegensatz zum Kopf kann der gesamte Kopf-Hals-Bereich nicht als einzelner starrer Körper betrachtet werden. Weicht die Position der Halswirbelsäule in der einen Aufnahme stark von der in der anderen Aufnahme ab, kann eine korrekte rigide Transformation nicht gefunden werden, weil die nicht-rigiden Bereiche einen negativen Einfluss auf das Bildähnlichkeitsmaß haben. Daher dürfen nicht-rigide Bereiche, also der Hals, nicht in die Berechnung des Bildähnlichkeitsmaßes einbezogen werden. Dies erfordert eine automatische grobe Detektion der Region-of-Interest (ROI), also des Kopfes.

Im Folgenden werden die in [14] veröffentlichten Heuristiken zur Detektion der ROI und zur Initialisierung vorgestellt, die eine vollautomatische rigide Registrierung unter Beachtung der gerade beschriebenen Anforderungen ermöglichen.

**Region-of-Interest-Detektion:** Zur Detektion des Übergangs zwischen Kopf und Hals wird die Tatsache ausgenutzt, dass die Fläche des Körpers in axialen Schichten des Halses kleiner ist als im Kopf. Dies gilt auch für abgeschnittene Aufnahmen. Unter der Annahme, dass der Hintergrund sowohl in CT- als auch in MRT-Aufnahmen niedrigere Intensitätswerte als der Körper aufweist, kann der Körper durch ein einfaches Schwellwertverfahren extrahiert werden. Um die Anfälligkeit für Rauschen gering zu halten, wird die Aufnahme zunächst mit einem Gaussfilter ( $\sigma = 3 \text{ mm}$ ) gefiltert. Der Schwellwert wird auf 10% des in der Aufnahme vorkommenden Intensitätsbereichs gesetzt. Zur Rauschunterdrückung wird anschließend die größte zusammenhängende Komponente



(a) Flächeninhalt des umschließenden Rechtecks in jeder Schicht ausgehend vom Hals (links) bis hin zum Kopf (rechts)



(b) Automatische Klassifikation in Kopf- (rot) und Hals-Voxel; Detektion des Kinns (blau) durch Finden des ventralsten, kaudalsten (grüne Linie) Voxel innerhalb des Kopfbereichs

Abbildung 4.21.: Automatische Detektion des Kopf-Hals-Übergangs

extrahiert, welche einer groben Maske des Kopf-Hals-Bereichs entspricht. In dieser Maske wird für jede axiale Schicht die maximale Ausdehnung in  $X$ - und  $Y$ -Richtung bestimmt und der Flächeninhalt des dadurch definierten Rechtecks berechnet. Auf den Verlauf des Flächeninhalts in  $Z$ -Richtung (siehe Abbildung 4.21(a)) wird zur Erhöhung der Robustheit ein Median-Filter mit einem Radius von  $15 \text{ mm}$  angewendet, bevor dieser zur Bestimmung des Kopf-Hals-Übergangs analysiert wird. Dazu wird von oben nach unten iteriert. Das erste lokale Maximum befindet sich sicherlich im Kopf-Bereich, das darauf folgende lokale Minimum befindet sich bereits im Hals. Von dort an wird nun in entgegengesetzte Richtung iteriert bis 105% des Flächeninhalts am lokalen Minimum überschritten wird. Die dazugehörige Schicht trennt nun die Kopf-Schichten (oberhalb) von den Hals-Schichten (unterhalb) (siehe Abbildung 4.21(b)).

**Initialisierung:** Zur Initialisierung ist die robuste Detektion einer gemeinsamen Landmarke erforderlich. Da die Heuristik auch auf abgeschnittene Bilder angewendet werden soll, kann die nächstliegende Landmarke – die Nasenspitze – nicht verwendet werden, da sie nicht notwendigerweise in allen Bildern zu sehen ist. Stattdessen wird die Kinnschmelze verwendet, welche zwar nicht sonderlich punktförmig ist, dafür aber in allen Aufnahmen, dessen Zielregion die Mundhöhle ist, zu sehen ist. Die Kinnschmelze ist durch den ventralsten (vordersten), kaudalsten (untersten) Voxel der durch den vorangegangenen Schritt bestimmten Maske gegeben (siehe Abbildung 4.21(b)). Für alle Voxel der Maske in den Kopf-Schichten wird dazu die Distanz zu einer Ebene berechnet, die diagonal zwischen die axiale und koronale Ebene vor dem Kopf platziert wird. Das gesuchte Voxel hat den minimalen Abstand zu dieser Ebene.

**Mutual-Information-basierte Registrierung:** Während in MRT-Aufnahmen das Weichteilgewebe auf einen großen Intensitätsbereich abgebildet wird, werden diese nur auf einen kleinen Bereich in CT-Aufnahmen (üblicherweise zwischen  $-125$  und  $225 \text{ HU}$ ) abgebildet. Da alle anderen Intensitäten, wie z.B. die vielen unterschiedlichen Grauwerte für Knochen, nicht sinnvoll für einen Vergleich

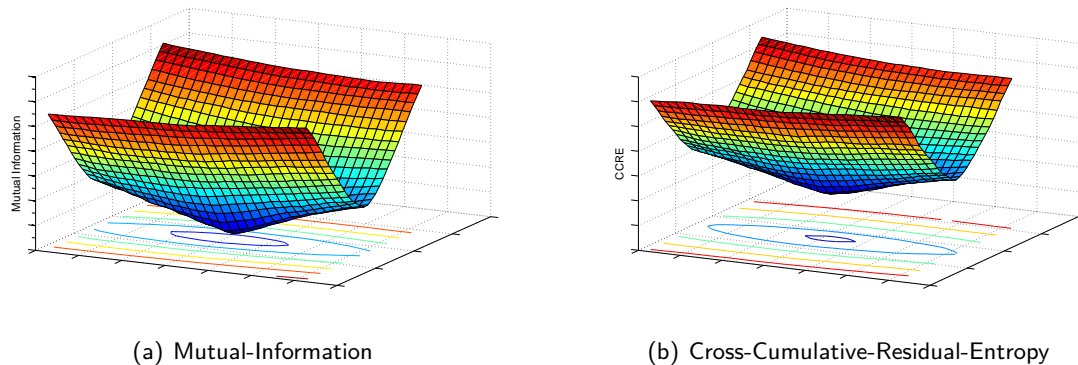


Abbildung 4.22.: Vergleich von *Mutual-Information* und *Cross-Cumulative-Residual-Entropy* bei einer  $30^\circ$  3D-Rotation um das Kinn. Das Optimum der dritten Dimension ist in die anderen beiden Dimensionen projiziert.

herangezogen werden können, werden alle Intensitäten der CT-Aufnahme, die sich außerhalb des Weichteilgewebefensters befinden, in einem Vorverarbeitungsschritt auf  $-125$  bzw. auf  $225$  abgebildet. Die MRT-Aufnahme – das Modellbild – wird dann initial so verschoben, dass die Landmarke – das Kinn – in beiden Aufnahmen übereinander liegt. Unter der Annahme, dass das Kinn nun perfekt ausgerichtet ist, reduziert sich das ursprünglich 6-dimensionale (Rotation + Translation) Registrierungsproblem auf ein 3-dimensionales Problem: Rotation um das Kinn. Diese Rotation wird durch die Maximierung einer geeigneten Ähnlichkeitsmetrik gefunden. Aufgrund der Multimodalität bieten sich statistische Maße, wie *Mutual-Information* oder *Cross-Cumulative-Residual-Entropy* [104] an. Wie in Abbildung 4.22 zu sehen ist, ist für dieses niedrigdimensionale Optimierungsproblem kein fundamentaler Unterschied zu erwarten. In der Implementierung wurde die Mutual-Information-Implementierung von Mattes et al. [114] verwendet, wobei in jedem Bild 50000 Samples zufällig ausgewählt wurden, um ein gemeinsames Histogramm der Dimension  $24 \times 24$  aufzubauen, aus dem die Einzelentropien und die gemeinsame Entropie berechnet werden. Um einen negativen Einfluss von potentiell deformiertem Gewebe – wie dem Hals – zu verhindern, findet die Auswahl der Samples nur in den zuvor als Kopf klassifizierten Schichten statt. Aufgrund der Glattheit und des hohen Konvergenzbereichs des Ähnlichkeitsmaßes (siehe Abbildung 4.22) konvergiert ein Gradientenabstiegsverfahren schnell gegen die benötigte Rotation.

Da das Kinn nicht punktförmig ist und die Detektion daher mittels der beschriebenen einfachen Heuristik ungenau ist, ist es nötig, etwaige Registrierungsungenauigkeiten in einem weiteren Schritt auszugleichen. Wiederum wird die Mutual-Information mit einem Gradientenabstiegsverfahren maximiert, diesmal allerdings in dem vollen rigid 6-dimensionalen Parameterraum. Um den Einfluss der Translation zu vergrößern, wird das Rotationszentrum diesmal vom Kinn in das geometrische Zentrum aller Kopf-Voxel verschoben. Das Resultat sind nun zwei Aufnahmen des Kopf-Hals-Bereichs, auf denen der Kopf bestmöglich ausgerichtet ist. Aufgrund möglicher Positionsunterschiede, insbesondere der Halswirbelsäule, kann aber nach wie vor ein großer Versatz außerhalb des Schädels vorliegen.

#### 4. Methoden

| Bez.               | Typ  | Beschreibung  |
|--------------------|--|---|
| $n$                | $\in \mathbb{N}$   | Anzahl der Atlaselemente  |
| $i$                | $\in [1, \dots, n]$  | Index eines Atlaselements   |
| $m$                | $\in \mathbb{N}$   | Anzahl der Trainingsdatensätze  |
| $j$                | $\in [1, \dots, m]$  | Index eines Trainingsdatensatzes  |
| $\mathcal{D}$      | $\subset \mathbb{R}^3$                                       | Definitionsbereich der Aufnahme   |
| $\mathcal{R}$      | $\subset \mathbb{R}$   | Grauwertbereich   |
| $\mathcal{C}^j$    | $\mathcal{D} \mapsto \mathcal{R}$                            | CT-Aufnahme des Trainingsdatensatzes $j$  |
| $\mathcal{S}_i^j$  | $\mathcal{D}_i \mapsto \{0, 1\}$                             | Segmentierung der Struktur $i$ im Trainingsdatensatz $j$  |
| $\mathcal{S}^j$    | $\mathcal{D} \mapsto \{0, 1\}$                               | Segmentierung aller Strukturen im Trainingsdatensatz $j$  |
| $T_i$              | $\mathbb{R}^7 \times \mathbb{R}^3 \mapsto \mathbb{R}^3$      | Ähnlichkeitstransformation des Elementkoordinatensystem von Element $i$ in das Atlaskoordinatensystem                               |
| $T_g$              | $\mathbb{R}^6 \times \mathbb{R}^3 \mapsto \mathbb{R}^3$      | Globale rigide Transformation des Atlaskoordinatensystems in das Weltkoordinatensystem  |
| $T^j$              | $\mathbb{R}^6 \times \mathbb{R}^3 \mapsto \mathbb{R}^3$      | Rigide Transformation des Atlaskoordinatensystems in das Koordinatensystem des Trainingsdatensatzes $j$                             |
| $T_i^j$            | $\mathbb{R}^7 \times \mathbb{R}^3 \mapsto \mathbb{R}^3$      | Ähnlichkeitstransformation des Elementkoordinatensystems von Element $i$ in das Koordinatensystem des Trainingsdatensatzes $j$      |
| $E_i^j$            | $\mathbb{R}^7 \times \mathbb{R}^3 \mapsto \mathbb{R}^3$      | Ähnlichkeitstransformation des Elementkoordinatensystems von Element $i$ in das Atlaskoordinatensystem des Trainingsdatensatzes $j$ |
| $\mathbf{p}_i$     | $\in \mathbb{R}^7$   | Parameter der Elementtransformation $T_i$   |
| $\mathbf{p}$       | $\in \mathbb{R}^{7n}$  | Parameter aller Elementtransformationen (Artikulationsparameter)  |
| $\mathbf{p}^j$     | $\in \mathbb{R}^{7n}$  | Artikulationsparameter des Trainingsdatensatzes $j$   |
| $\mathbf{e}$       | $\in \mathbb{R}^6$   | Parameter der globalen Transformation $T_g$   |
| $\mathcal{D}_i$    | $\subset \mathbb{R}^3$                                       | Definitionsbereich des Atlaselements $i$  |
| $\mathcal{W}_i$    | $\mathbb{R}^3 \mapsto [0, 1]$                                | Wahrscheinlichkeitsbild des Atlaselements $i$ zur Formrepräsentation  |
| $\mathcal{A}_i$    | $\mathcal{D}_i \mapsto \mathcal{R}$                          | Texturrepräsentation des Atlaselements $i$  |
| $\odot$            | $\mathbb{R} \times \mathcal{T} \mapsto \mathcal{T}$          | Skalarprodukt einer Transformation  |
| $\oplus$           | $\mathcal{T} \times \mathcal{T} \mapsto \mathcal{T}$         | Kommutative Addition zweier Transformationen  |
| $k$                | $\in \mathbb{N}$   | Dimensionalität des gelernten Artikulationsraums  |
| $\mathbf{a}$       | $\in \mathbb{R}^k$   | Parameter einer Artikulation im gelernten Artikulationsraum   |
| $\mathbf{r}$       | $\in \mathbb{R}^{7n}$  | Fehlervektor einer Artikulation zum gelernten Artikulationsraum   |
| $P$                | $\in \mathbb{R}^{7n \times m}$                               | Parameter aller Trainingsdatensätze   |
| $\bar{\mathbf{p}}$ | $\in \mathbb{R}^{7n}$  | Parameter der Durchschnittsartikulation   |
| $A$                | $\in \mathbb{R}^{7n \times k}$                               | Basisvektoren des gelernten Artikulationsraums  |
| $\lambda_j$        | $\in \mathbb{R}$   | Eigenwerte der Kovarianzmatrix der Trainingsparameter   |
| $E_{\text{Bild}}$  | $\mathbb{R}^{7n} \times \mathbb{R}^6 \mapsto \mathbb{R}_0^+$ | Bildenergieterm zur Anpassung des artikulierten Atlas   |
| $E_{\text{Art.}}$  | $\mathbb{R}^{7n} \mapsto \mathbb{R}_0^+$                     | Artikulationsrestriktionsterm zur Erhaltung der gelernten Artikulation  |
| $\psi$             | $\in \mathbb{R}^+$   | Balance zwischen Bild- und Artikulationsenergieterm   |
| $\rho$             | $\in \mathbb{R}^+$   | Balance zwischen Fehler innerhalb des und zum Artikulationsraum   |
| $t$                | $\in \mathbb{R}^+$   | Anpassungsfortschritt   |
| $w_i$              | $\mathbb{R}^+ \mapsto [0, 1]$                                | Bildenergietermeinfluss des Atlaselements $i$ in Anhängigkeit von $t$   |
| $\mathcal{L}$      | $\mathcal{D} \mapsto \{0, \dots, n\}$                        | Labelbild; Zuordnung jedes Voxels zu einem Atlaselement   |

Tabelle 4.13.: Übersicht der Notation des artikulierten Atlas

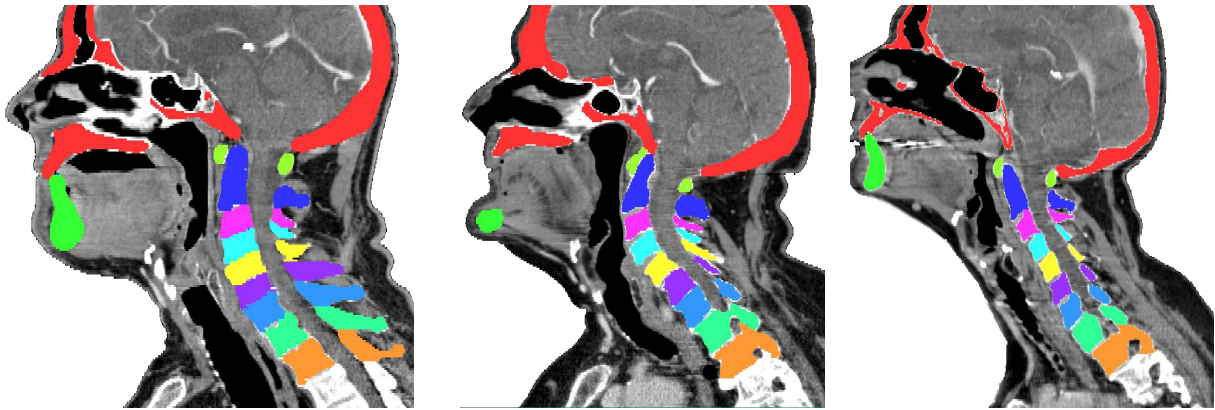


Abbildung 4.23.: Beispieltrainingsdatensätze zum Aufbau des artikulierten Atlas. Jedes manuell segmentierte Objekt ist in einer anderen Farbe dargestellt.

#### 4.2.2. Artikulierter Atlas zur Segmentierung einzelner Knochen

Aufgrund des hohen Kontrastes kann Knochen sehr einfach, bspw. durch ein Schwellwertverfahren, in CT-Aufnahmen segmentiert werden. In dem Szenario dieser Arbeit ist es allerdings nicht ausreichend, Knochen per se zu segmentieren, sondern jeder Knochen des Kopf-Hals-Bereichs, der eine gewisse Größe überschreitet, soll getrennt segmentiert werden. Dabei handelt es sich um den Schädel, den Unterkiefer, die 7 Halswirbel (C1-C7) und die ersten beiden Brustwirbel (Th1, Th2). Da die Formen, insbesondere die der Wirbel, recht komplex sind und diese sich auch noch in unmittelbarer Umgebung voneinander befinden, ist die Segmentierung nicht trivial. Globales Formwissen muss eingebracht werden. Doch auch das allein ist nicht ausreichend, weil benachbarte Strukturen sehr ähnliche Charakteristiken aufweisen und so ein möglicher Formmodell-basierter Segmentierungsansatz sich potentiell an benachbarte Strukturen anpassen könnte. Der Ausweg, der in dieser Arbeit gegangen wird, besteht darin, alle Knochen gleichzeitig mithilfe eines neuartigen *Artikulierten Atlas* zu segmentieren, welcher in [15] veröffentlicht wurde und neben Wissen über Textur und Form der einzelnen Strukturen auch die Lagebeziehungen der Strukturen untereinander statistisch gelernt hat. Dieser Abschnitt beschäftigt sich mit dem Aufbau dieses Atlas. Im darauf folgenden Abschnitt wird dann die eigentliche Segmentierung durch die Anpassung des Atlas an einen ungesesehenen Datensatz beschrieben. Die in diesen Abschnitten verwendete Notation ist in Tabelle 4.13 zusammengefasst.

**Trainingsgrundlage:** Trainingsgrundlage für den Atlas stellen manuell annotierte CT-Datensätze des Kopf-Hals Bereichs dar. Jeder der insgesamt  $m$  Trainingsdatensätze verfügt über eine CT-Aufnahme:  $\mathcal{C}^j : \mathcal{D} \mapsto \mathcal{R}, j \in \{1, \dots, m\}$ .  $\mathcal{D} \subset \mathbb{R}^3$  ist dabei der Definitionsbereich der Aufnahme und  $\mathcal{R} \subset \mathbb{R}$  der zulässige Grauwertbereich. In dieser CT-Aufnahme sind alle  $n$  Strukturen segmentiert. Diese Segmentierungen werden als Binärbilder dargestellt:  $\mathcal{S}_i^j : \mathcal{D} \mapsto \{0, 1\}, i \in \{1, \dots, n\}$ .

Da die schichtweise manuelle Segmentierung von vielen Strukturen sehr zeitaufwändig ist (ca. 3 Stunden pro Datensatz durch einen geübten Anwender), wurde der konkret in dieser Arbeit verwendete Atlas auf nur 15 unterschiedlichen Datensätzen trainiert. Die Auswahl erfolgte dabei so, dass eine möglichst große Variabilität an Form und vor allem an Wirbelsäulenpositionen abgedeckt wird. Da der Kopf-Hals-Bereich des Menschen rechts-links symmetrisch ist, wurden alle Datensätze und da-



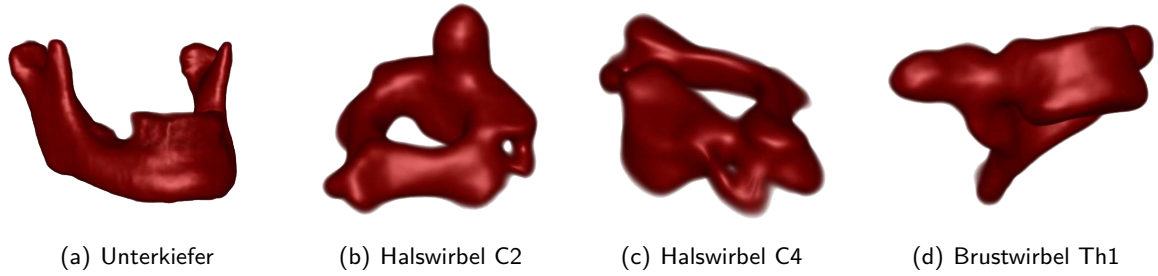


Abbildung 4.24.: Die Formvarianz der einzelnen Atlaselemente wird durch Wahrscheinlichkeitsbilder repräsentiert. Einige Beispiele sind hier durch Volumenrendering abgebildet.

zugehörige Segmentierungen zusätzlich an der sagittalen Ebene gespiegelt und den Trainingsdaten hinzugefügt. Dadurch wird ein größerer Bereich an Form und Objektlagen abgedeckt. Die Anzahl der Trainingsdatensätze beträgt also  $m = 30$ . Einige Beispiele sind in Abbildung 4.23 dargestellt.

**Aufbau:** Der artikulierte Atlas besteht aus  $n$  Elementen. Jedem dieser Elemente ist genau eine zu segmentierende Struktur (also bspw. der Schädel oder der Wirbel C4) zugeordnet.

Jedes Element verfügt über ein(e)

- **Transformation:** Die Transformation  $\mathcal{T}_i : \mathbb{R}^7 \times \mathbb{R}^3 \mapsto \mathbb{R}^3$  transformiert einen Punkt  $\mathbf{x} \in \mathbb{R}^3$  im Elementkoordinatensystem durch einen 7-dimensionalen Parametervektor  $\mathbf{p}_i \in \mathbb{R}^7$  in einen Punkt  $\mathbf{y} \in \mathbb{R}^3$  im Atlaskoordinatensystem. Die sieben Freiheitsgrade setzen sich aus einer Translation, einer 3D-Rotation und einer isotropen Skalierung zusammen. Auf weitere Freiheitsgrade wurde aus Komplexitätsgründen verzichtet. Auf eine weitere Einschränkung der Freiheitsgrade, bspw. durch eine explizite Modellierung von Gelenken, wie dies in der Literatur (z.B. [125]) geschieht, soll an dieser Stelle verzichtet werden, um zunächst beliebige Konstellationen der Atlaselemente zu ermöglichen. Die Einschränkung auf plausible Konstellationen erfolgt erst später durch das Lernen des Artikulationsraums aus Trainingsdaten.
- **Form:** Da die Genauigkeit der Segmentierung in dem Szenario dieser Arbeit eine untergeordnete Rolle spielt, kann auf eine Modellierung der Oberfläche durch ein statistisches Formmodell [45][145] verzichtet werden. Stattdessen gibt ein Wahrscheinlichkeitsbild  $\mathcal{W}_i : \mathcal{D}_i \subset \mathbb{R}^3 \mapsto [0, 1]$  statistischen Aufschluss darüber, ob ein Punkt sich innerhalb oder außerhalb des Elements befindet. Beispiele dafür sind in Abbildung 4.24 dargestellt.
- **Textur:** Die Textur (das Aussehen) eines Elements in einer CT-Aufnahme wird durch ein Durchschnittsbild  $\mathcal{A}_i : \mathcal{D}_i \mapsto \mathcal{R}$  repräsentiert, welches aus den Trainings-CT-Datensätzen erzeugt wird. Abbildung 4.25 zeigt einige Beispiele.

Der gesamte artikulierte Atlas befindet sich in einem Atlaskoordinatensystem. Dieses ist durch eine globale rigide Transformation  $\mathcal{T}_g : \mathbb{R}^6 \times \mathbb{R}^3 \mapsto \mathbb{R}^3$  gegeben, welche einen Punkt im Atlaskoordinatensystem in einen Punkt im Weltkoordinatensystem transformiert. Punkte, welche im Elementkoordinatensystem vorliegen, werden also durch die Konkatenation der Elementtransformation und der globalen Transformation  $\mathcal{T}_g \circ \mathcal{T}_i$  in das Weltkoordinatensystem transformiert.



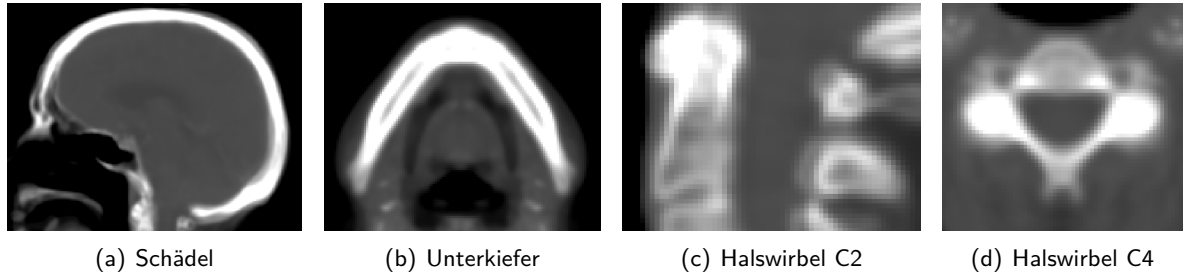


Abbildung 4.25.: Durchschnitintensitätsbilder modellieren die Textur der Atlaselemente. Ausgewählte Schichten sind hier abgebildet.

**Parametrisierung:** Da jedes der  $n$  Atlaselemente 7 Freiheitsgrade hat, kann die *Artikulation*, also die Bewegung bzw. Pose des Atlas, durch einen  $7n$ -dimensionalen internen Parametervektor  $\mathbf{p} \in \mathbb{R}^{7n}$  beschrieben werden. Zusätzlich wird die globale Position und Ausrichtung des Atlas durch einen 6-dimensionalen externen Parametervektor  $\mathbf{e} \in \mathbb{R}^6$  beschrieben, welcher den Parametern der globalen rigiden Transformation entspricht.

Allerdings bleibt bei dieser Beschreibung zunächst unklar, wie die 7 Parameter eines Elements nun die Elementtransformation parametrisieren sollen. Um die Anforderungen an eine geeignete Parametrisierung formulieren zu können, wird ein Blick auf die beiden Anwendungsszenarien geworfen:

1. Beim *Artikulationstraining* soll der Raum aller in der Praxis auftretenden Transformationen aus einzelnen Beispieltransformationen  $\mathcal{T}^j$  gelernt werden. Neben den Trainingsbeispielen sollen auch alle 'dazwischenliegenden' Transformationen in diesem Raum enthalten sein. Dazu ist es nötig, zwischen den Transformationen stetig interpolieren zu können.
2. Bei der späteren *Anpassung* des Atlas an einen Datensatz soll der trainierte Artikulationsraum berücksichtigt werden. Je weiter eine Transformation  $\mathcal{T}$  von dem gelernten Raum 'entfernt' ist, desto unwahrscheinlicher soll  $\mathcal{T}$  sein. Dazu ist es nötig, einen Abstand zwischen zwei Transformationen zu definieren.

Beide Anforderungen sind in einem *Vektorraum* erfüllt. Linearkombinationen zweier Elemente können durch Skalarmultiplikation und Addition gebildet werden und durch eine Norm kann eine Abstandsmetrik induziert werden.

Bei der Elementtransformation handelt es sich um eine sog. Ähnlichkeitstransformation. Streckenverhältnisse und Winkelgrößen bleiben unverändert, Streckenlängen ändern sich jedoch im Allgemeinen. Eine solche Transformation wird oftmals durch einen Skalierungsfaktor  $s \in \mathbb{R}$  eine Rotationsmatrix  $R \in \mathbb{R}^{3 \times 3}$  und einen Translationsvektor  $\mathbf{t} \in \mathbb{R}^3$  beschrieben:

$$\mathbf{y} = T(\mathbf{x}) = s \cdot R\mathbf{x} + \mathbf{t} \quad (4.60)$$

Rotationsmatrizen besitzen 3 Freiheitsgrade und werden oft entweder durch *Eulerwinkel* oder durch einen *Einheitsquaternion* erzeugt. Die Transformation kann also durch einen 7-dimensionalen Vektor beschrieben werden. Unglücklicherweise erfüllt diese Art der Parametrisierung nicht die genannten Anforderungen: Bspw. entspricht der Nullvektor nicht dem neutralen Element, d.h. der Identität, und auch die Addition zweier solcher Vektoren führt nicht, wie intuitiv erwartet, zu einer Konkatenation der Transformationen.

Um diesem Problem zu entgegnen, schlug Alexa [146] eine Parametrisierung von Transformationen auf Basis der Matrixexponentialfunktion vor, die es erlaubt, auf Transformationen Operationen durchzuführen, so wie dies in einem Vektorraum geschieht. Eine Transformation kann in homogenen Koordinaten durch eine  $4 \times 4$ -Matrix  $T$  beschrieben werden:

$$T = \begin{pmatrix} & sR & \mathbf{t} & \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (4.61)$$

Diese wird durch die Matrixexponentialfunktion einer weiteren  $4 \times 4$ -Matrix  $A$  beschrieben:

$$T = e^A, \quad (4.62)$$

wobei die Matrixexponentialfunktion analog zur skalaren Exponentialfunktion durch die Reihe  $e^{(A)} = \sum_{k=0}^{\infty} \frac{A^k}{k!}$  definiert ist und effizient berechnet werden kann (siehe [146]). Durch die Umkehrfunktion, also den Matrixlogarithmus, kann die Parametrisierung  $A = \log(T)$  berechnet werden. In dieser Darstellung ist durch  $r \odot T = e^{r \log(T)}$  ein Skalarprodukt  $\odot$  und durch  $X \oplus Y = e^{\log(X) + \log(Y)}$  eine Addition  $\oplus$  definiert, welche intuitiv erwartete Eigenschaften erfüllen:

- Die Skalarmultiplikation einer beliebigen Transformation mit 0 ergibt die Identität ( $0 \odot X = I$ ).
- Die Skalarmultiplikation einer beliebigen Transformation mit  $-1$  ergibt die inverse Transformation ( $-1 \odot X = X^{-1}$ ).
- Die Addition ist kommutativ ( $X \oplus Y = Y \oplus X$ ).
- Die Addition zweier Hälften einer Transformation ergibt die Transformation selbst ( $\frac{1}{2}X \oplus \frac{1}{2}X = X$ ).
- Die Addition zweier Rotationen ergibt wieder eine Rotation.

Die Darstellung einer Transformation durch das Matrixexponential scheint also für die beiden genannten Anwendungsgebiete geeignet zu sein. Jedoch können in dieser Darstellung alle möglichen Transformationen und nicht nur Ähnlichkeitstransformationen dargestellt werden. Die Anzahl der Freiheitsgrade muss also reduziert werden. Da eine Addition und eine Skalarmultiplikation vorliegt, kann dies durch Linearkombinationen von  $k$  Basistransformationen  $B_i$  geschehen:

$$T(\mathbf{p}) = e^{\left( \sum_{i=1}^k p_i \log(B_i) \right)}, \quad (4.63)$$

wobei  $\mathbf{p}$  der  $k$ -dimensionale Parametervektor ist, der die einzelnen Basistransformationen gewichtet. Für den konkreten Fall der Ähnlichkeitstransformation werden also  $k = 7$  Basistransformationen benötigt: 3 Translation, 3 Rotation und eine Skalierung. Diese wurden ad-hoc auf Translation um je 1 mm in x-, y- und z-Richtung, Rotation je um  $2^\circ$  um x-, y- und z-Achse sowie Skalierung um 2% festgelegt. Jede beliebige Ähnlichkeitstransformation lässt sich daraus kombinieren. Mithilfe der Pseudoinversen lassen sich aus einer beliebigen Ähnlichkeitstransformation  $T$  die Parameter  $\mathbf{p}$  bestimmen.

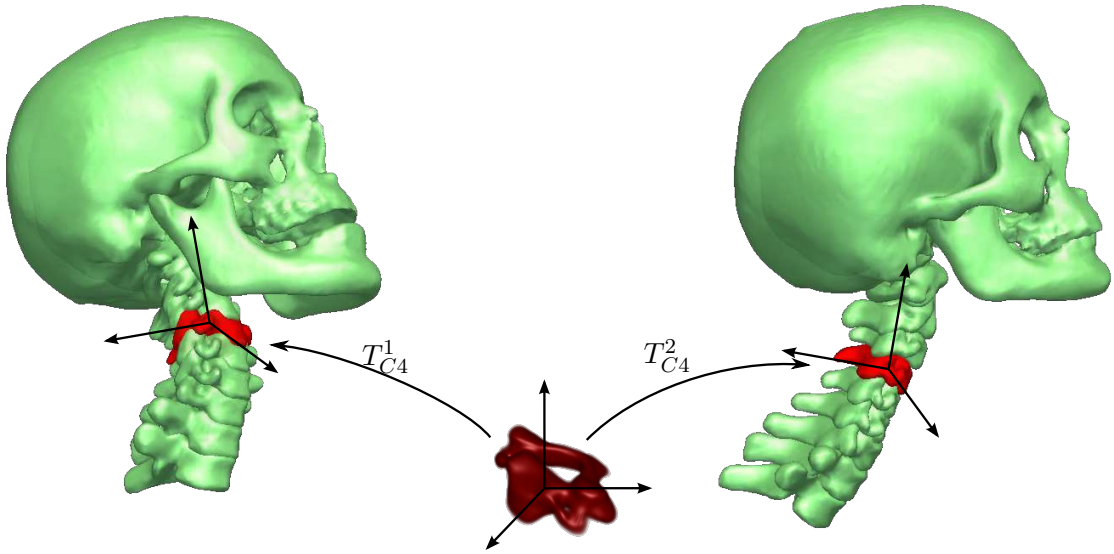


Abbildung 4.26.: Die Transformationen  $T_i^j$  überführen das Koordinatensystem des Wahrscheinlichkeitsbildes  $\mathcal{W}_i$  von Element  $i$  in das Koordinatensystem der Trainingsdaten  $j$ .

**Formtraining:** Wie bereits erwähnt, wird die Form eines Elements  $i$  durch ein Wahrscheinlichkeitsbild  $\mathcal{W}_i$  repräsentiert. Dieses wird durch Mittelung der Segmentierungen der Trainingsdaten  $\mathcal{S}_i^j$  berechnet:

$$\mathcal{W}_i(\mathbf{x}) = \frac{1}{m} \sum_{j=1}^m \mathcal{S}_i^j(T_i^j(\mathbf{x})) , \quad (4.64)$$

wobei  $T_i^j$  Ähnlichkeitstransformationen sind, welche die einzelnen Segmentierungen in ein gemeinsames Koordinatensystem überführen (siehe Abbildung 4.26). Die Bestimmung der Transformationen  $T_i^j$  und die Berechnung des Wahrscheinlichkeitsbildes geschieht iterativ durch folgenden Algorithmus:

1. Setze  $\mathcal{S}_i^1$  als Referenzbild  $\mathcal{R}_i$ .
2. Verschiebe den Schwerpunkt von  $\mathcal{R}_i$  in den Ursprung.
3. Finde für alle Trainingssegmentierungen  $\mathcal{S}_i^j$  je eine Transformation  $T_i^j$  zum Referenzbild durch Bildregistrierung:
  - a) Initialisiere die Transformation  $T_i^j$  durch eine Verschiebung des Schwerpunkts von  $\mathcal{R}_i$  in den Schwerpunkt von  $\mathcal{S}_i^j$ .
  - b) Finde durch ein Gradientenabstiegsverfahren diejenige Ähnlichkeitstransformation  $T_i^j$ , welche den mittleren quadratischen Fehler vom transformierten  $\mathcal{R}_i$  zu  $\mathcal{S}_i^j$  minimiert.
4. Berechne das Wahrscheinlichkeitsbild  $\mathcal{W}_i$  mithilfe Gleichung 4.64 aus allen Trainingsdaten
5. Berechne die Durchschnittstransformation  $\bar{T}_i$  aller  $T_i^j$  mithilfe der Matrixexponentialdarstellung.

6. Wende die inverse Durchschnittstransformation  $(\bar{T}_i)^{-1}$  auf  $\mathcal{W}_i$  an.
7. Verschiebe den Schwerpunkt von  $\mathcal{W}_i$  in den Ursprung.
8. Glätte  $\mathcal{W}_i$  durch einen Gaussfilter.
9. Terminiere falls  $\mathcal{W}_i \approx \mathcal{R}_i$  gilt, ansonsten setze  $\mathcal{R}_i = \mathcal{W}_i$  und gehe zu Schritt 3.

Durch diesen iterativen Prozess wird der Einfluss der Trainingsdatenreihenfolge minimal gehalten. Als Definitionsbereich  $\mathcal{D}_i$  des Wahrscheinlichkeitsbildes wird ein Quader gewählt, welcher alle Orte  $\mathbf{x}$  mit der Eigenschaft  $\mathcal{W}_i(\mathbf{x}) > 0$  enthält.

**Training der Textur:** Die Textur eines Elements  $i$  ist durch ein Durchschnittsbild  $\mathcal{A}_i$  repräsentiert, welches mithilfe der zuvor bestimmten Transformationen  $T_i^j$  aus den CT-Aufnahmen  $\mathcal{C}^j$  der Trainingsdatensätze  $j$  berechnet wird:

$$\mathcal{A}_i(\mathbf{x}) = \frac{1}{m} \sum_{j=1}^m \mathcal{C}^j \left( T_i^j(\mathbf{x}) \right) . \quad (4.65)$$

Die Transformationen stellen sicher, dass sich bei der Mittelung alle CT-Aufnahmen im selben Koordinatensystem, und zwar in dem des betrachteten Elements, befinden. An Orten mit hoher Formvarianz oder Orten, die weit vom Element entfernt sind, ist das Durchschnittsbild deutlich unschärfer als an Orten im Elementzentrum mit geringer Formvarianz. Der Definitionsbereich  $\mathcal{D}_i$  von  $\mathcal{A}_i$  entspricht dem des Wahrscheinlichkeitsbildes  $\mathcal{W}_i$  zur Formrepräsentation. Diese Art der Texturrepräsentation wird später im Anpassungsprozess verwendet, um eine Aussage über die Überlappung des artikulierten Atlas und einer CT-Aufnahme treffen zu können.

**Artikulationstraining:** Durch das Artikulationstraining sollen die relativen Lagen der Atlaselemente zueinander gelernt werden. Da dies unabhängig von der absoluten Position und Ausrichtung der Trainingsdaten geschehen soll, müssen diese erst in ein gemeinsames Koordinatensystem überführt werden. Dies geschieht durch denselben Algorithmus, der schon beim Formtraining zur Bestimmung der Transformationen der einzelnen Elemente angewendet wurde. Als Eingabedaten dienen diesmal allerdings die Vereinigungen der Segmentierungen aller Elemente eines Trainingsdatensatzes  $\mathcal{S}^j = \bigcup_{\forall i} \mathcal{S}_i^j$ . Statt Ähnlichkeitstransformationen werden diesmal jedoch rigide Transformationen  $T^j$  gesucht, weil die Größe der Trainingsdaten in das Artikulationstraining einbezogen werden soll. Die Transformationen  $T^j$  bestimmen das Atlaskoordinatensystem für jeden Trainingsdatensatz. Das entstandene Wahrscheinlichkeitsbild ist nur ein Beiprodukt und wird im Folgenden nicht mehr benötigt.

Aus den absoluten Transformationen  $T^j$  der Trainingsdatensätze sowie den absoluten Transformationen  $T_i^j$  der Modellelemente der Trainingsdatensätze können nun relative Elementtransformationen  $E_i^j$  berechnet werden (siehe Abbildung 4.27):

$$E_i^j = T_i^j \circ (T^j)^{-1} . \quad (4.66)$$

Da es sich bei den Elementtransformationen  $E_i^j$  um Ähnlichkeitstransformationen handelt, können diese durch je einen 7-dimensionalen Parametervektor  $\mathbf{p}_i^j \in \mathbb{R}^7$  beschrieben werden. Dazu kommt die zuvor beschriebene Parametrisierung in Matrixexponentialdarstellung zum Einsatz. Der Vektor  $\mathbf{p}_i^j$  bildet die Gewichte zur Linearkombination von Basistransformationen.

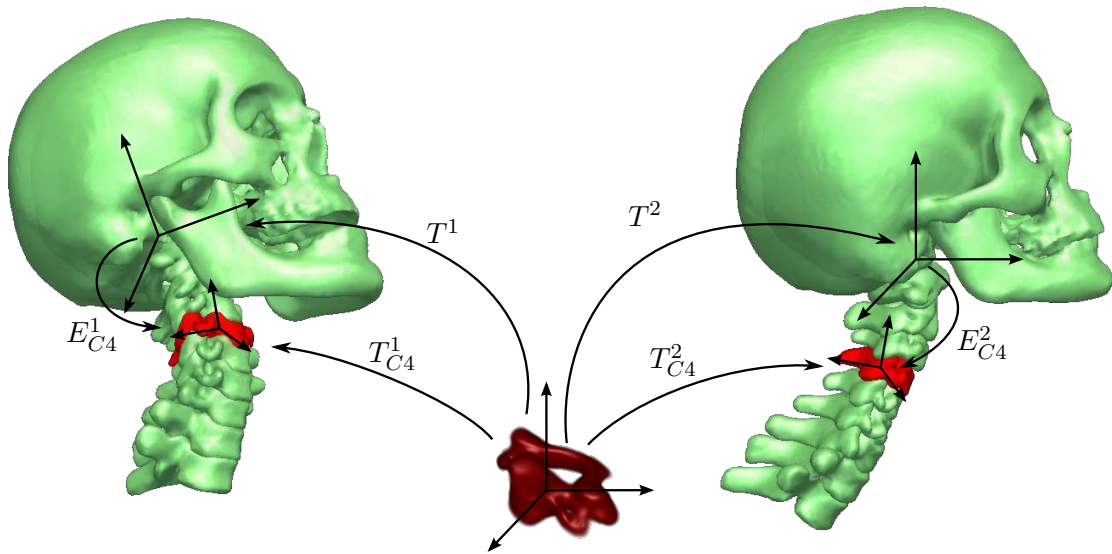


Abbildung 4.27.: Die Trainingsdatentransformation  $T^j$  überführt das Atlaskoordinatensystem in das Trainingsdatenkoordinatensystem. Die Elementtransformationen  $E_i^j$  befinden sich dann für alle Trainingsdaten im selben Koordinatensystem und die Artikulation kann daraus gelernt werden.

Um die gesamte Artikulation eines Trainingsdatensatzes zu beschreiben, werden die Parametervektoren  $\mathbf{p}_i^j$  der Elemente zu einem einzigen  $7n$ -dimensionalen Parametervektor  $\mathbf{p}^j \in \mathbb{R}^{7n}$  zusammengefasst. Aus diesen  $m$  Parametervektoren  $\mathbf{p}^j$  soll nun der Raum aller typischen relativen Lagen von Modellelementen gelernt werden.

Wie bereits motiviert, kann durch die Parametrisierung der Transformationen mithilfe des Matrix-exponentials mit den Vektoren  $\mathbf{p}^j$  wie in einem Vektorraum gerechnet werden. Daher kann auch das Standardverfahren zum Lernen von Formmodellen, die *Hauptkomponentenanalyse* oder englisch *Principal Component Analysis* (PCA) [147], verwendet werden. Die Parametervektoren der Trainingsdatensätze werden dazu in eine  $7n \times m$ -Matrix  $P = (\mathbf{p}^1, \dots, \mathbf{p}^m)$  geschrieben. Eine beliebige Artikulation  $\mathbf{p} \in \mathbb{R}^{7n}$  wird durch einen  $k \ll 7n$ -dimensionalen Artikulationsvektor  $\mathbf{a} \in \mathbb{R}^k$  dargestellt:

$$\mathbf{p} = \bar{\mathbf{p}} + A \cdot \mathbf{a} + \mathbf{r}, \quad (4.67)$$

wobei  $\bar{\mathbf{p}} = \frac{1}{m} \sum_{j=1}^m \mathbf{p}^j$  der Durchschnittsartikulation aller Trainingsartikulationen entspricht. Die Matrix  $A$  enthält die  $k$  Eigenvektoren mit den größten zugehörigen Eigenwerte  $\lambda_j, \lambda_1 \geq \dots \geq \lambda_k \geq \dots \geq \lambda_{m-1} > 0$  der Kovarianzmatrix  $C = \frac{1}{m-1} \hat{P} \hat{P}^T$ ,  $\hat{P} = (\mathbf{p}_1 - \bar{\mathbf{p}}, \dots, \mathbf{p}_m - \bar{\mathbf{p}})$ . Der durch die  $k$  Eigenvektoren aufgespannte Raum wird als *Artikulationsraum* bzw. Featurespace ( $F$ ) bezeichnet. Alle darin enthaltenen Artikulationen können durch  $\mathbf{a}$  beschrieben werden. Der erste Eintrag von  $\mathbf{a}$  beschreibt die Bewegung entlang der größten Varianz in den Trainingsdaten, der zweite Eintrag die Bewegung entlang der zweitgrößten Varianz und so weiter. Da die Daten starke Korrelationen aufweisen (bspw. liegt der Halswirbel C4 immer knapp unterhalb von C3), kann davon ausgegangen werden, dass sehr wenige Freiheitsgrade ausreichen, um die Trainingsdaten mit ausreichender

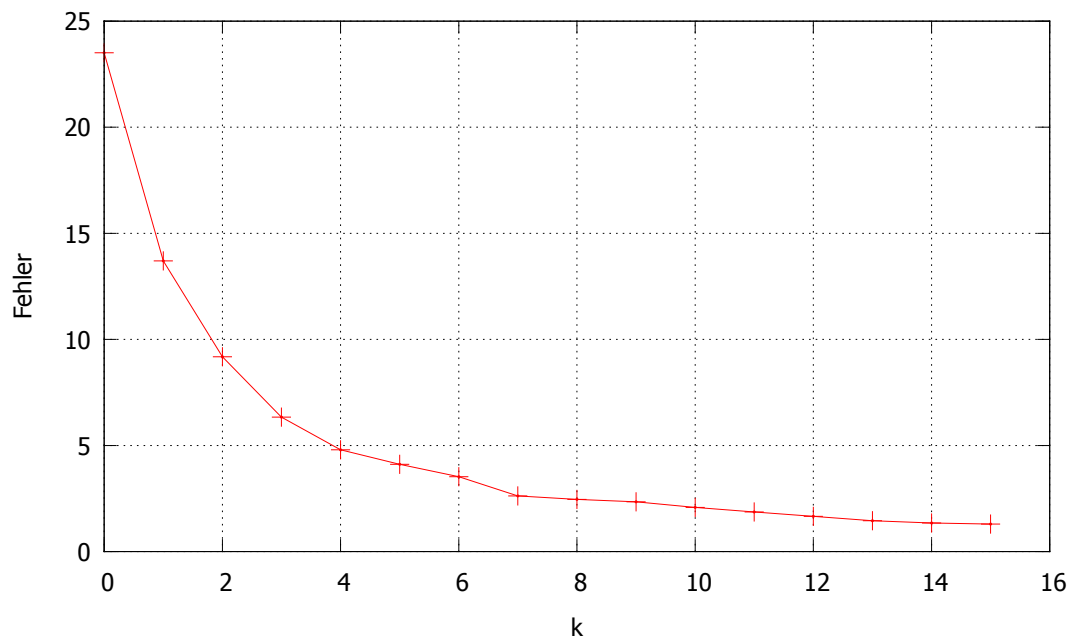


Abbildung 4.28.: Länge des Fehlervektors bei unterschiedlicher Dimensionalität  $k$  des Artikulationsvektors  $\mathbf{a}$

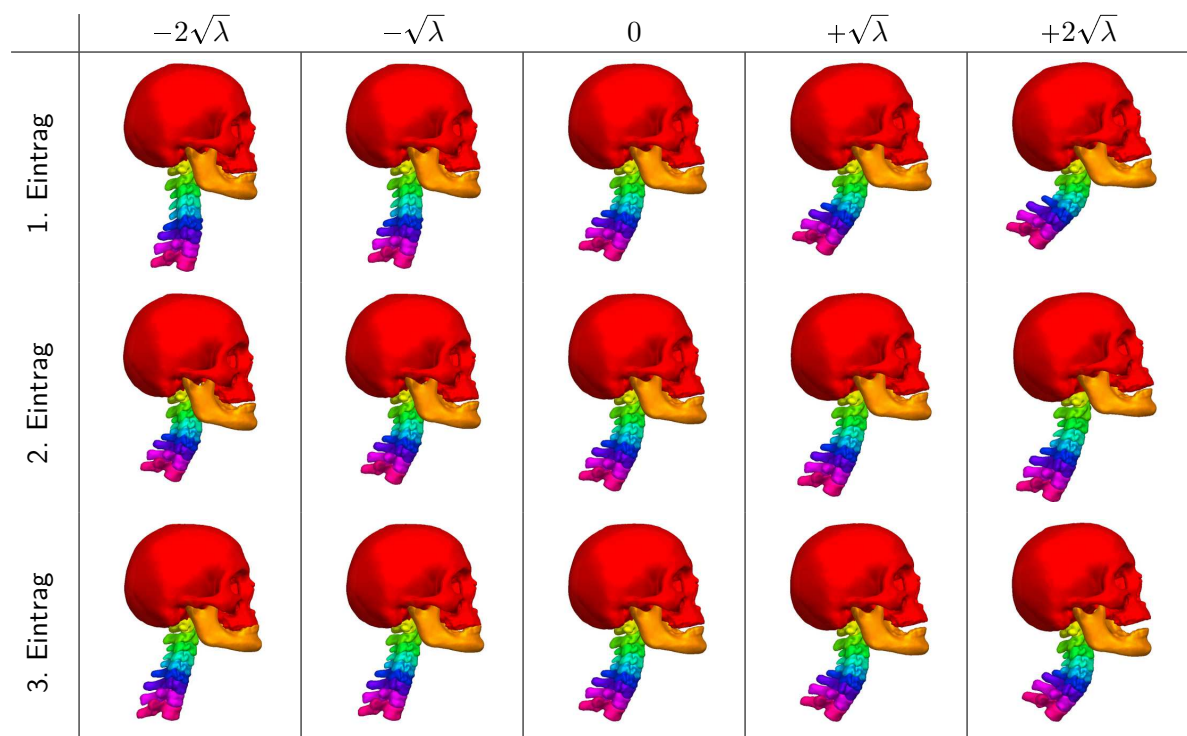


Abbildung 4.29.: Die Artikulation des Atlas in Abhängigkeit von den ersten drei Einträgen des Artikulationsvektors  $\mathbf{a}$ . Die Durchschnittsartikulation  $\bar{\mathbf{p}}$  ist in der mittleren Spalte dargestellt.

Genauigkeit zu beschreiben. Alle nicht durch  $\mathbf{a}$  beschreibbaren Artikulationen befinden sich in dem komplementären Raum  $\bar{F}$ . Sie werden durch den Residualvektor  $\mathbf{r} \in \mathbb{R}^{7n}$  beschrieben. Da jedoch ein Großteil der in der Praxis vorkommenden Artikulationen innerhalb oder zumindest nahe an  $F$  sein sollte, weist der Residualvektor  $\mathbf{r}$  üblicherweise eine geringe Länge auf.

Es soll nun geklärt werden, wie viele Freiheitsgrade tatsächlich nötig sind. Dazu wurde für unterschiedliche  $k$  eine Leave-One-Out-Kreuzvalidierung angewendet. Der Artikulationsraum wurde für  $m - 1$  Parametervektoren gelernt, der übriggebliebene Parametervektor wurde mithilfe von Gleichung 4.67 rekonstruiert und schließlich wurde die Länge des Residualvektors  $\|\mathbf{r}\|$  berechnet. Dies wurde für alle Kombinationsmöglichkeiten von  $m - 1$  Parametervektoren durchgeführt und die entstandenen Fehler wurden gemittelt. Abbildung 4.28 zeigt diesen mittleren Fehler für unterschiedliche Dimensionalität  $k$  des Artikulationsvektors  $\mathbf{a}$ . Wie zu erwarten, nimmt der Fehler mit steigendem  $k$  zunächst rapide ab, für höhere  $k$  jedoch nicht mehr so stark und ab ca.  $k = 10$  fast gar nicht mehr. Daher wurde  $k = 10$  festgelegt. Der ursprünglich 77-dimensionale Raum kann also aufgrund der vorhandenen Korrelation durch einen 10-dimensionalen Vektor effizient beschrieben werden. Abbildung 4.29 zeigt die Artikulation in Abhängigkeit der ersten drei Einträge.

### 4.2.3. Anpassung des artikulierten Atlas

Um einen ungesehenen CT-Datensatz mithilfe des artikulierten Atlas segmentieren zu können, muss dieser an den Datensatz angepasst werden. Dies geschieht durch die Minimierung eines Energiefunktional der Form

$$E(\mathbf{p}, \mathbf{e}) = E_{\text{Bild}}(\mathbf{p}, \mathbf{e}) + \psi E_{\text{Artikulation}}(\mathbf{p}), \quad (4.68)$$

wobei  $\mathbf{p}$  die Artikulation und  $\mathbf{e}$  die globale Position und Orientierung beschreibt. Der Bildenergieterm  $E_{\text{Bild}}$  liefert umso geringere Werte, desto mehr die Grauwerte im CT-Datensatz an den durch  $\mathbf{p}$  und  $\mathbf{e}$  gegebenen Positionen mit der gelernten Textur übereinstimmen. Der Artikulationsrestriktionsterm liefert hingegen umso geringere Werte, desto mehr die durch  $\mathbf{p}$  beschriebene Artikulation der des zuvor gelernten Artikulationsraums entspricht. Letzterer ist unabhängig von der globalen Position und Orientierung  $\mathbf{e}$  des Atlas. Der Parameter  $\psi$  balanciert die beiden Terme.

Diese Art von Energiefunktional kam schon an einigen Stellen in dieser Arbeit zum Einsatz, jedes Mal jedoch mit einer leicht anderen Bedeutung. In Abschnitt 2.2.1 wurden Segmentierungsverfahren mit lokalem Formwissen eingeführt. Dazu wird in Gleichung 2.1 Bildinformation und lokale Formhaltung balanciert. Dasselbe geschieht in Abschnitt 3.1.5, 4.1.2 und 4.1.4 zur Segmentierung von Lymphknoten. Darüber hinaus wird in Abschnitt 2.2.2.1 die Gleichung 2.6 dazu verwendet, um bei der bildbasierten deformierbaren Registrierung eine Balance zwischen der Übereinstimmung des Referenzbildes mit dem deformierten Modellbild und lokaler Plausibilität des Deformationsfeldes zu finden.

**Bildenergieterm:** Da CT-Aufnahmen durch die Hounsfield-Skala normiert sind und sonstige Intensitätsschwankungen bspw. durch Artefakte als vernachlässigbar angenommen werden, wird der mittlere gewichtete quadratische Fehler zwischen der transformierten Atlastextur  $\mathcal{A}_i$  und der CT-

Aufnahme  $\mathcal{C}$  als Fehlermaß verwendet. Für das Atlaselement  $i$  ergibt sich der Fehler zu

$$E_{\text{Bild}_i}(\mathbf{p}_i, \mathbf{e}) = \frac{\sum_{l=1}^{m_i} G_\sigma \star \mathcal{W}_i(\mathbf{x}_i^l) [\mathcal{A}_i(\mathbf{x}_i^l) - \mathcal{C}(T(\mathbf{e}) \circ T(\mathbf{p}_i) \circ \mathbf{x}_i^l)]^2}{\sum_{l=1}^{m_i} G_\sigma \star \mathcal{W}_i(\mathbf{x}_i^l)}, \quad (4.69)$$

wobei  $\mathbf{p}_i$  der Anteil der Gesamtartikulation  $\mathbf{p}$  ist, welcher die Transformation für das Element  $i$  parametrisiert. Der mittlere quadratische Fehler wird durch eine Menge, bestehend aus  $m_i$  zufällig aus dem Elementdefinitionsbereich  $\mathcal{D}_i$  gezogenen Stichproben (Samples)  $\mathbf{x}_i^l, l \in [1, \dots, m_i]$  approximiert.  $m_i$  wird dabei linear in Abhängigkeit von der Gesamtgröße gewählt. Je größer  $m_i$  ist, desto genauer ist die Approximation. Gleichzeitig erhöht sich aber auch die Laufzeit proportional. 20000 Samples für alle 11 Atlaselemente haben sich als guter Kompromiss zwischen Laufzeit und Approximationsgüte erwiesen. Die Gewichtung des Fehlers erfolgt durch das Gauss-gefilterte ( $G_\sigma$ ) Wahrscheinlichkeitsbild  $\mathcal{W}_i$ . Dies hat zur Folge, dass Bereiche, welche wahrscheinlicher zu dem Element gehören oder - durch die Filterung - sich zumindest in der Nähe des Elements befinden, stärker gewichtet werden als die anderen Bereiche. So ein Vorgehen ist sinnvoll, weil die Grauwerte außerhalb des Elements einer größeren Varianz unterliegen (vgl. Training der Textur in Abschnitt 4.2.2).

Die Gesamtbildenergie ergibt sich dann als gewichtete Summe der Elementenergien:

$$E_{\text{Bild}}(\mathbf{p}, \mathbf{e}) = \sum_{i=1}^n w_i E_{\text{Bild}_i}(\mathbf{p}_i, \mathbf{e}), \quad (4.70)$$

wobei die Gewichte  $w_i$  so gewählt sind, dass sie proportional zur Größe des Elements sind ( $w_i \sim \int_{\mathbf{x} \in \mathcal{D}_i} \mathcal{W}_i(\mathbf{x}) d\mathbf{x}$ ) und  $\sum_{i=1}^n w_i = 1$  gilt. Auf diese Weise ist sichergestellt, dass kleine Elemente (bspw. ein Wirbel) einen kleineren Einfluss auf die Gesamtenergie als große Elemente (bspw. der Schädel) haben.

**Artikulationsrestriktionsterm:** Die mögliche Artikulation soll sich auf den gelernten Artikulationsraum beschränken oder sich zumindest in der Nähe des Raums befinden. Modelliert man die Wahrscheinlichkeitsdichte der Artikulation näherungsweise als Gaußverteilung, kann das Fehlermaß nach Moghaddam und Pentland [148] als Artikulationsrestriktionsterm verwendet werden:

$$E_{\text{Artikulation}}(\mathbf{p}) = \sum_{l=1}^k \frac{a_l^2(\mathbf{p})}{\lambda_l} + \frac{1}{\rho} \|\mathbf{r}(\mathbf{p})\|^2, \quad (4.71)$$

wobei  $\lambda_l, \lambda_1 \geq \dots \geq \lambda_{m-1} > 0$  die Eigenwerte der Kovarianzmatrix  $C$  sind (siehe Artikulationstraining in Abschnitt 4.2.2).

Die Beschreibung des Anteils der Artikulation innerhalb des durch die Eigenvektoren von  $C$  aufgespannten Raums (gelernter Artikulationsraum  $F$ ) folgt aus Gleichung 4.67 und geschieht durch den Vektor  $\mathbf{a} = A^T \cdot (\mathbf{p} - \bar{\mathbf{p}})$ . Der dazugehörige Fehlerterm (der erste Term von Gleichung 4.71) ist durch Mahalanobis-Distanz gegeben. Je größer eine Koordinate relativ zu der erwarteten Varianz – gegeben durch  $\lambda$  – ist, desto größer ist der Fehler. Dieser Fehler wird als *Distance In Feature Space* (DIFS) bezeichnet.



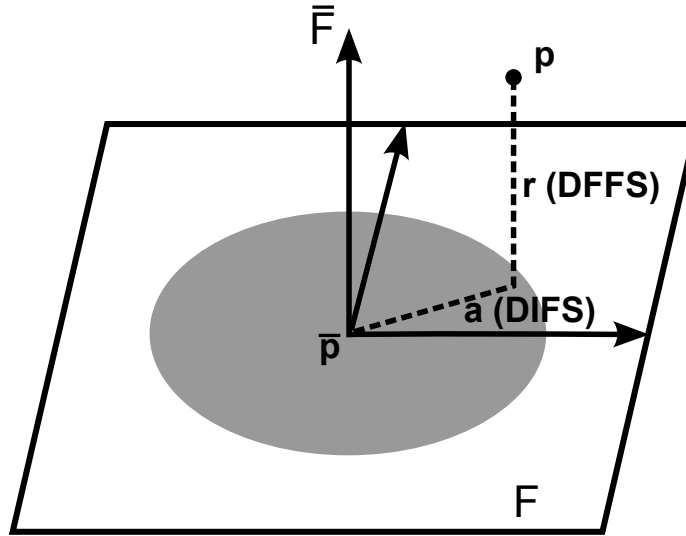


Abbildung 4.30.: Zusammenhang zwischen dem Fehler innerhalb des Artikulationsraums  $F$  und im komplementären Raum  $\bar{F}$

Da sich eine Artikulation  $\mathbf{p}$  nicht notwendigerweise innerhalb des gelernten Artikulationsraums  $F$  befindet,  $\mathbf{p}$  aber zumindest in der Nähe von  $F$  sein sollte, entspricht der zweite Term der Gleichung 4.71 der quadratischen euklidischen Distanz von  $\mathbf{p}$  zu  $F$ . Dies ist der Fehler in dem komplementären Raum  $\bar{F}$  und wird als *Distance From Feature Space* (DFFS) bezeichnet. Abbildung 4.30 veranschaulicht den Zusammenhang zwischen DIFS und DFFS.

Der Parameter  $\rho$  wird als DFFS-Koeffizient bezeichnet und balanciert beide Fehlerquellen. Unter der Annahme einer Gaußverteilung kann dieser nach [148] optimal aus den Trainingsdaten bestimmt werden:

$$\rho = \frac{1}{m - 1 - k} \sum_{j=k+1}^{m-1} \lambda_j . \quad (4.72)$$

Ist die Varianz in  $\bar{F}$  hoch, erhält man große Werte für  $\rho$  und der Einfluss des DFFS auf den Artikulationsrestriktionsterm ist gering.

**Energieminimierung durch Gradientenabstieg:** Die Anpassung des Atlas an einen Datensatz  $\mathcal{C}$  wird durch Minimierung des in Gleichung 4.68 beschriebenen Energiefunktional  $E(\mathbf{p}, \mathbf{e})$  realisiert. Die Minimierung läuft in einem zweistufigen Prozess ab:

1. In einem ersten Schritt wird die optimale globale Position und Orientierung  $\mathbf{e}^*$  gefunden. Dies geschieht durch ein Gradientenabstiegsverfahren mit konstanter Schrittlänge.

Die Initialisierung geschieht durch eine Heuristik: Da CT-Aufnahmen üblicherweise dieselbe Ausrichtung besitzen, wird initial keine Rotation angenommen. Die initiale Position wird auf das kraniale (obere) Ende der CT-Aufnahme gesetzt, weil oberhalb des Kopfes keine weiteren Strukturen erwartet werden, an die sich der Atlas fälschlicherweise anpasst. Diese Heuristik setzt voraus, dass ein gewisser Teil des Schädels in der CT-Aufnahme sichtbar ist. Es hat sich herausgestellt, dass in etwa die Hälfte des Schädels ausreichend ist.

Der Artikulationsvektor wird während dieses Schrittes auf die Durchschnittsartikulation  $\mathbf{p} = \bar{\mathbf{p}}$  gesetzt. Der Artikulationsrestriktionsterm hat immer den Wert 0 und muss darum nicht berechnet werden.

2. In einem zweiten Schritt wird die optimale Artikulation  $\mathbf{p}^*$  gemeinsam mit der optimalen globalen Position und Orientierung  $\mathbf{e}^*$  wiederum durch ein lokales Gradientenabstiegsverfahren gefunden. Dies geschieht ausgehend von der Durchschnittsartikulation in zuvor bestimmter globaler Position und Orientierung. Da  $\mathbf{e}^*$  in dem Schritt zuvor für die Durchschnittsartikulation bestimmt wurde und diese für abweichende  $\mathbf{p}$  ebenfalls anders sein kann, kann  $\mathbf{e}$  in diesem Schritt nicht konstant gelassen werden.

Für das Gradientenabstiegsverfahren müssen die partiellen Ableitungen  $\frac{\partial E(\mathbf{p}, \mathbf{e})}{\partial e_i}$  und  $\frac{\partial E(\mathbf{p}, \mathbf{e})}{\partial p_i}$  berechnet werden. Dies kann weitestgehend analytisch geschehen:

$$\nabla E(\mathbf{p}, \mathbf{e}) = \nabla E_{\text{Bild}}(\mathbf{p}, \mathbf{e}) + \psi \nabla E_{\text{Artikulation}}(\mathbf{p}) , \quad (4.73)$$

wobei für die Einträge  $\frac{\partial E_{\text{Artikulation}}}{\partial e_i} = 0$  gilt, weil der Artikulationsrestriktionsterm nicht von der globalen Position und Orientierung abhängt.

Zur Berechnung des Gradienten des Artikulationsrestriktionsterms wird unter Ausnutzen von

$$\|\mathbf{r}\|^2 = \|\mathbf{p} - \bar{\mathbf{p}}\|^2 - \|\mathbf{a}\|^2 , \quad (4.74)$$

was aus dem Satz des Pythagoras folgt (siehe Abbildung 4.30), wird Gleichung 4.71 zunächst umgeschrieben:

$$\begin{aligned} E_{\text{Artikulation}}(\mathbf{p}) &= (\mathbf{p} - \bar{\mathbf{p}})^T A \Lambda^{-1} A^T (\mathbf{p} - \bar{\mathbf{p}}) \\ &+ \frac{1}{\rho} [(\mathbf{p} - \bar{\mathbf{p}})^T (\mathbf{p} - \bar{\mathbf{p}}) - (\mathbf{p} - \bar{\mathbf{p}})^T A A^T (\mathbf{p} - \bar{\mathbf{p}})] , \end{aligned} \quad (4.75)$$

wobei  $\Lambda^{-1} = \text{diag}(\frac{1}{\lambda_1}, \dots, \frac{1}{\lambda_k})$  gilt. Gleichung 4.74 gilt, weil das Residual  $\mathbf{r}$  stets orthogonal zu dem durch  $F$  aufgespannten Raum ist und somit der Satz des Pythagoras gilt (siehe Abbildung 4.30). Der Gradient des Artikulationsrestriktionsterms kann dann durch Vektoranalysis aus Gleichung 4.75 bestimmt werden:

$$\nabla E_{\text{Artikulation}}(\mathbf{p}) = 2A\Lambda^{-1}A^T(\mathbf{p} - \bar{\mathbf{p}}) + \frac{2}{\rho}(\mathbf{p} - \bar{\mathbf{p}} - AA^T(\mathbf{p} - \bar{\mathbf{p}})) . \quad (4.76)$$

Die Berechnung des Gradienten des Bildenergieterms gestaltet sich etwas komplizierter. Für ein Atlaselement  $i$  und dessen Parameter  $\mathbf{p}_i$  ergibt sich dieser zu:

$$\nabla E_{\text{Bild}_i}(\mathbf{p}_i) = \frac{\sum_{j=1}^{m_i} G_\sigma \star \mathcal{W}_i(\mathbf{x}_i^j) 2 \left[ \mathcal{A}_i(\mathbf{x}_i^j) - \mathcal{C}(\cdot) \right] [\nabla(-\mathcal{C}(\cdot))] \circ T(\mathbf{e}) \circ J_{T(\mathbf{p}_i)} \circ \mathbf{x}_i^j}{\sum_{j=1}^{m_i} G_\sigma \star \mathcal{W}_i(\mathbf{x}_i^j)} , \quad (4.77)$$

wobei  $\mathcal{C}(\cdot) = \mathcal{C}(T(\mathbf{e}) \circ T(\mathbf{p}_i) \circ \mathbf{x}_i^j)$  gilt und  $J_{T(\mathbf{p}_i)}$  die Jacobimatrix der durch  $\mathbf{p}_i$  parametrisierten Transformation ist. Die Berechnung dieser Jacobimatrix basiert auf der Ableitung des Matrixexponentials. Dies kann entweder durch Auswertung des in [149] beschriebenen Integrals erfolgen oder durch finite Differenzen approximiert werden. In dieser Arbeit kam aus Effizienzgründen Letzteres

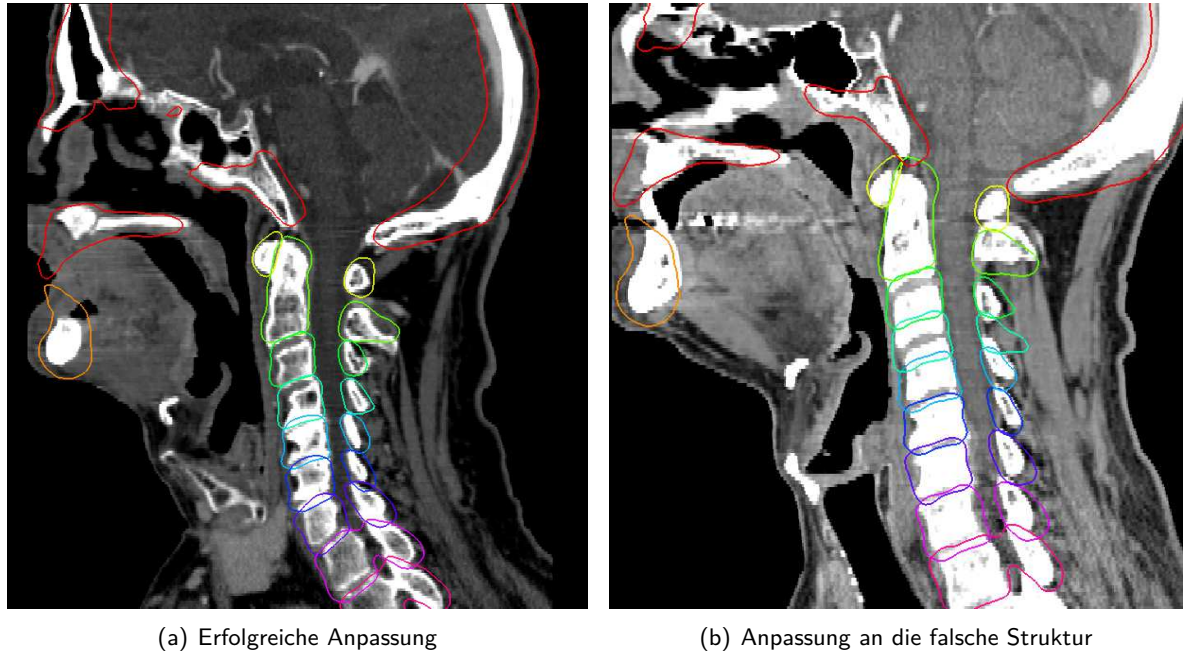


Abbildung 4.31.: Anpassung des artikulierten Atlas an unterschiedliche Datensätze durch das Gradientenabstiegsverfahren

zur Anwendung. Die Ableitung nach den Parametern der globalen Position und Orientierung erfolgt sehr ähnlich, wobei hierbei jeder Summand aus Gleichung 4.70 einen Einfluss auf den Gradienten hat und somit berücksichtigt werden muss.

Basierend auf dem gerade beschriebenen Gradienten kann nun die Aktualisierungsregel für den  $k$ -ten Schritt des Gradientenabstiegsverfahrens formuliert werden:

$$(\mathbf{p}, \mathbf{e})^{k+1} = (\mathbf{p}, \mathbf{e})^k - \Delta \cdot \nabla E(\mathbf{p}, \mathbf{e}) , \quad (4.78)$$

wobei  $\Delta$  der Schrittweite entspricht. Beendet wird die Anpassung, sobald  $(\mathbf{p}, \mathbf{e})^{k+1} \approx (\mathbf{p}, \mathbf{e})^k$  gilt und der artikuliert Atlas sich somit in einem lokalen Minimum befindet.

Abbildung 4.31 zeigt Beispiele für den angepassten artikulierten Atlas. Während mit dieser Anpassungsstrategie in einigen Fällen gute Ergebnisse erzielt werden können, kann es in anderen Fällen passieren, dass das Verfahren in das falsche lokale Energieminimum konvergiert. Gerade innerhalb der Wirbelsäule tritt dies auf, wie das Beispiel in Abbildung 4.31(b) zeigt. Der Grund dafür ist, dass benachbarte Wirbel sehr ähnlich aussehen. Somit kommt es sehr stark auf die Initialisierung an, ob bspw. das Modellelement, welches den Wirbel C5 repräsentiert, tatsächlich in den Wirbel C5 oder doch in den benachbarten Wirbel C4 oder C6 konvergiert. Bei einzelnen Wirbeln kann dies durch den Artikulationsrestriktionsterm kompensiert werden. Bewegt sich jedoch eine genügend große Anzahl an Elementen in Richtung der falschen Struktur, kann dies nicht mehr kompensiert werden und zwei Atlaselemente befinden sich auf derselben Struktur oder ein Atlaselement deckt zwei benachbarte Strukturen ab. Dieser Nachteil des puren Gradientenabstiegsverfahrens tritt insbesondere bei Datensätzen auf, welche untypisch sind, also stark von der Durchschnittsartikulation  $\bar{\mathbf{p}}$  abweichen. Ein Beispiel dafür sind Patienten, deren Anatomie durch Pathologien oder chirurgische Eingriffe von der

normalen Anatomie abweicht. Sowohl das Abdecken aller möglichen Anatomieabweichungen während des Artikulationstrainings als auch die Verwendung eines globalen Optimierungsverfahrens ist durch mangelnde Trainingsdaten bzw. aus Komplexitätsgründen in der Praxis nicht realisierbar.

**Energieminimierung durch adaptiven Gradientenabstieg:** Eine neuartige Lösung zum gerade beschriebenen Problem soll nun in diesem Abschnitt vorgestellt werden. Die Grundidee besteht darin, die Anpassung zunächst mit eindeutigen Strukturen zu beginnen und dann Nachbarschaftsbeziehungen im weiteren Anpassungsprozess zu berücksichtigen.

Umgesetzt ist diese Idee durch das Einführen von adaptiven Gewichten des Bildenergieterms. Gleichung 4.70 wird folgendermaßen erweitert:

$$E_{\text{Bild}}(\mathbf{p}, \mathbf{e}, t) = \sum_{i=1}^n w_i(t) E_{\text{Bild}_i}(\mathbf{p}_i, \mathbf{e}) , \quad (4.79)$$

wobei  $t \in [0, \dots, 1]$  ein neuer Parameter ist, welcher den Fortschritt des Anpassungsprozesses beschreibt und die Gewichte  $w_i(t)$  der Elemente nun davon abhängen. Für jedes Element wird nun ein Zeitpunkt  $t_i$  festgelegt, ab dem  $w_i(t) = w_i$  gelten soll. Vor diesem Zeitpunkt gilt  $w_i(t) = 0$  und die Anpassung des Elements geschieht ausschließlich aufgrund des Artikulationsrestriktionsterms. Während des Anpassungsprozesses mithilfe von Gleichung 4.78 wird nun auch der Parameter  $t$  ausgehend von  $t = 0$  in jedem Schritt leicht erhöht. Ist  $t = 1$  erreicht, ist die zu minimierende Energie zu der in Gleichung 4.68 formulierten Energie identisch.

Für das Kopf-Hals-Modell bedeutet dies konkret, dass die Anpassung mit dem Schädel beginnt ( $t_{\text{Schädel}} = 0$ ), sich dann über den Unterkiefer und den ersten Wirbel ( $t_{\text{Unterkiefer}} = 0.1$  und  $t_{\text{C1}} = 0.1$ ) bis hin zum untersten Wirbel fortpflanzt ( $t_{\text{Th2}} = 0.9$ ). Der Schädel dient damit als eindeutige Anfangsstruktur, welche zuerst gefunden wird. Ausgehend davon werden dann die Nachbarstrukturen gefunden, dann deren Nachbarn, usw. Durch die passive Anpassung mithilfe des Artikulationsrestriktionsterms ist sichergestellt, dass sobald ein Element  $i$  aktiv wird, d.h.  $t = t_i$  gilt, dieses sich bereits an einer gelernten plausiblen Stelle befindet. Die Anpassung an eine falsche Struktur wird dadurch unwahrscheinlich. Durch dieses sequenzielle Vorgehen wird verhindert, dass schlecht initialisierte Elemente die korrekte Anpassung des gesamten Atlas gefährden.

Abbildung 4.32 zeigt die Entwicklung der adaptiven Gewichte  $w_i(t)$  im Laufe eines Anpassungsprozesses. Anstatt eine harte Änderung von  $w_i(t) = 0, t < t_i$  auf  $w_i(t) = w_i, t \geq t_i$  zu verwenden, geschieht dies hier jedoch kontinuierlich mit Hilfe der Gaußfunktion. Die Anpassung von  $t$  erfolgt aufgrund der Konvergenz einzelner Atlaselemente. Ist bspw. der Betrag der partiellen Ableitungen der Parameter für den Schädel klein, ist dieser also gut angepasst, dann wird  $t$  erhöht und der Unterkiefer und das obere Ende der Wirbelsäule können sich nun an den Datensatz anpassen. Diese kontinuierliche Anpassung beschleunigt den Anpassungsprozess im Vergleich zu einer harten Änderung. So muss ein Element nicht *warten* bis das Vorgängerelement sich komplett angepasst hat.

**Segmentierung:** Sobald der artikulierte Atlas an einen Datensatz angepasst ist, also  $\mathbf{e}$  und  $\mathbf{p}$  bekannt sind, kann jedem Voxel in dem Datensatz ein *Label* durch eine Abbildung  $\mathcal{L} : \mathcal{D} \mapsto \{0, \dots, n\}$  zugeordnet werden. Die Labels entsprechen dabei dem Index des Atlaselements bzw. dem ausgezeichneten Label 0, falls das Voxel zu keinem Element zugeordnet werden soll. Eine einfache Übertragung

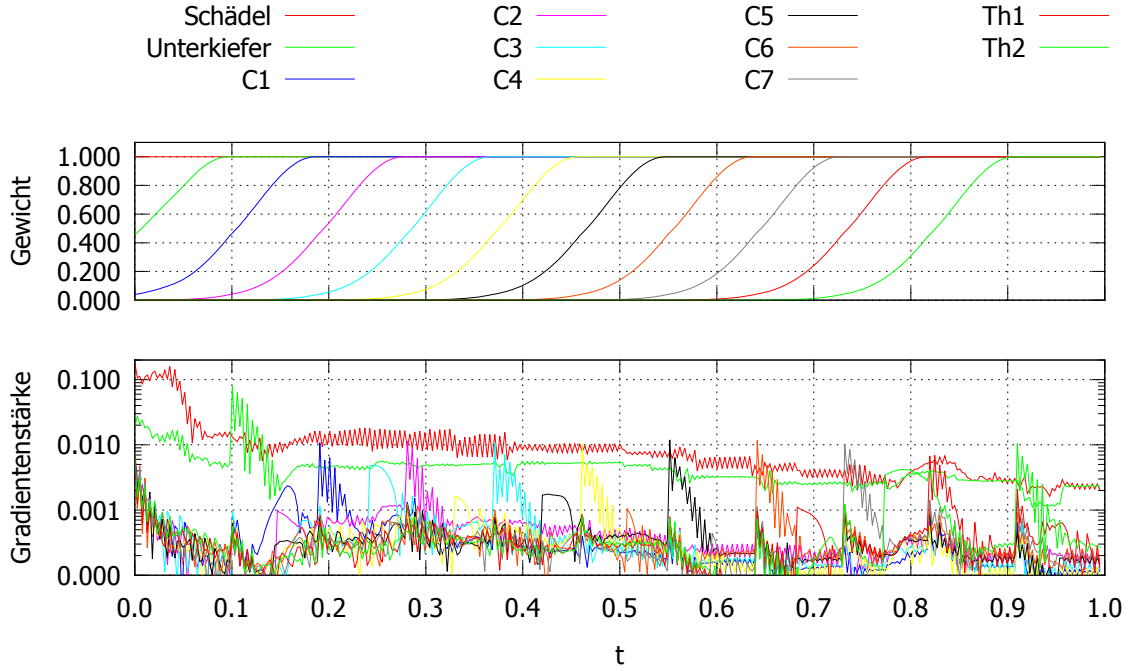


Abbildung 4.32.: Entwicklung der adaptiven Gewichte  $w_i(t)$  und der Energiegradientenstärke  $\nabla E_i$  für die verschiedenen Atlaselemente im Verlauf des Anpassungsprozesses

der Labels aus dem Atlas, wie dies üblicherweise bei atlasbasierter Segmentierung geschieht, ist hier nicht möglich, weil der artikulierte Atlas die Forminformation über Wahrscheinlichkeiten kodiert.

Stattdessen bildet die Grundlage für die Zuordnung zum einen die gelernte Information über die Form der zu segmentierenden Strukturen durch die Wahrscheinlichkeitsbilder  $\mathcal{W}_i$  und zum anderen Wissen über die Intensität von Knochen in CT-Aufnahmen. Aus den Trainingsdatensätzen wurde mithilfe des Bayestheorems die bedingte Wahrscheinlichkeit  $p(\text{Knochen}|I) = \frac{p(I|\text{Knochen})p(\text{Knochen})}{p(I)}$  für das Vorhandensein von Knochen bei einer gegebenen Intensität  $I$  numerisch bestimmt. Wie in Abbildung 4.33 zu sehen ist, kann diese gut durch die Arkustangensfunktion modelliert werden. Die kleine, aber dennoch positive Wahrscheinlichkeit für Knochen bei sehr geringen Bildintensitäten, sogar unter 0 HU, erscheint zunächst unplausibel, kann jedoch durch den Partialvolumeneffekt erklärt werden.

Die Wahrscheinlichkeit  $p_i(\mathbf{x})$  eines Elements  $i$  für den Ort eines Voxels  $\mathbf{x}$  ergibt sich zu

$$p_i(\mathbf{x}) = p(\text{Knochen}|\mathcal{C}(\mathbf{x})) \cdot \mathcal{W}_i(T^{-1}(\mathbf{p}_i) \circ T^{-1}(\mathbf{e}) \circ \mathbf{x}) . \quad (4.80)$$

Ausgehend davon wird die Labelzuordnung bestimmt:

$$\mathcal{L}(\mathbf{x}) = \begin{cases} 0 & ; \max_{\forall i} (p_i(\mathbf{x})) < p_{\min} \\ \arg\max_{\forall i} p_i(\mathbf{x}) & ; \text{sonst} \end{cases} , \quad (4.81)$$

wobei  $p_{\min} = 0.05$  ein Schwellwert für eine valide Elementzuordnung ist. Es wird also immer dasjenige Element ausgewählt, welches die höchste Wahrscheinlichkeit aufgrund der Form besitzt. Der Schwellwert  $p_{\min}$  sorgt dafür, dass Orte, welche nicht in der gelernten Form enthalten sind oder aber keine

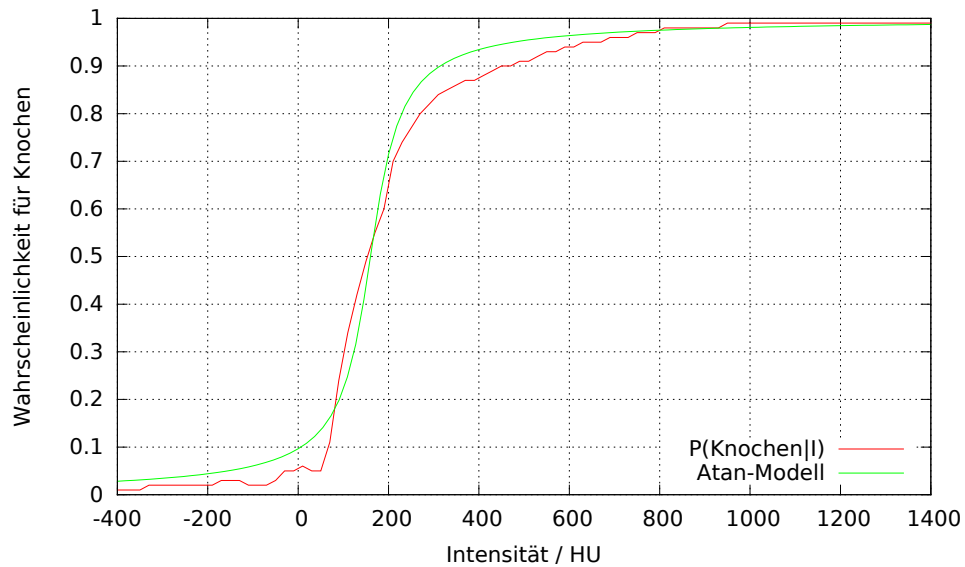


Abbildung 4.33.: Die bedingte Wahrscheinlichkeit für das Vorhandensein von Knochen abhängig von der CT-Intensität kann gut durch die Arkustangensfunktion modelliert werden.

knochentypischen CT-Intensitäten aufweisen, nicht zu einem Element gezählt werden, sondern als Hintergrund angesehen werden. Aus dem Labelbild können nun die Segmentierungen  $\mathcal{S}_i : \mathcal{D} \mapsto \{0, 1\}$  der Atlaselemente extrahiert werden.

### 4.2.4. Sukzessive lokal rigide Registrierung einzelner Knochen

Nun soll ausgehend von den Segmentierungen  $\mathcal{S}_i$  der CT-Aufnahme für jedes Element eine rigide Transformation  $\mathcal{T}_i$  zum Modellbild gefunden werden. Unter der Annahme, dass die einzelnen Knochen sich nicht oder nur geringfügig über die Aufnahmen hinweg ändern, ist eine rigide Transformation vollkommen ausreichend, weil keine Deformation, sondern nur eine andere Position und Ausrichtung erwartet wird. Diese Annahme wird dadurch unterstützt, dass Knochen deutlich härter als das umliegende Weichteilgewebe ist. Einzig und allein bei einer Anatomieänderung durch einen chirurgischen Eingriff ist es möglich, dass diese Annahme verletzt wird und somit die Registrierung möglicherweise fehlschlägt.

Die lokal rigide Registrierung eines Elements geschieht rein bildbasiert. Ein Bildähnlichkeitsmaß wird als Funktion der Transformationsparameter aufgefasst und mithilfe eines Gradientenabstiegsverfahrens wird ein lokales Maximum gefunden.

Das Maß wird dabei nur in einer geringfügig dilatierten Version von  $\mathcal{S}_i$  berechnet. Wird die Dilatation zu klein gewählt, ist nicht genügend Kontext für eine erfolgreiche Registrierung vorhanden, wird sie allerdings zu groß gewählt, hat möglicherweise deformiertes Weichteilgewebe einen negativen Einfluss auf das Registrierungsergebnis. In der Praxis hat sich eine Dilatation um 5 mm als geeignet erwiesen.

Als Bildähnlichkeitsmaß kommt, wie auch schon zur rigiden Registrierung des Kopfes (siehe Abschnitt 4.2.1), Mutual-Information [72][73] zum Einsatz. Mit ihr können sowohl bei monomodalen

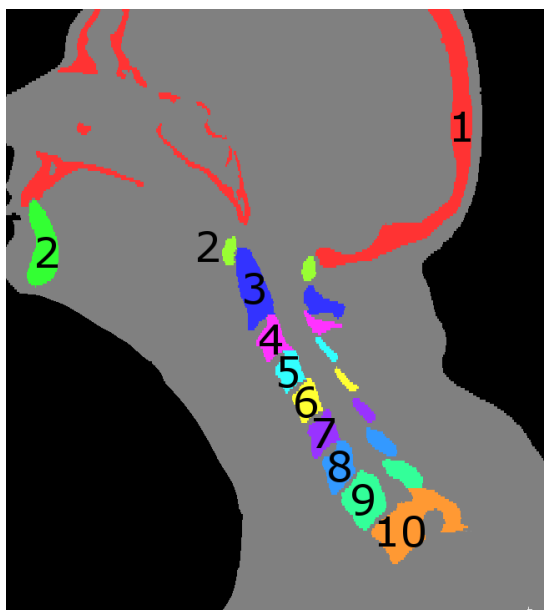


Abbildung 4.34.: Die sukzessive lokal rigide Registrierung in der dargestellten Reihenfolge ermöglicht es, auch sehr weit voneinander entfernte Strukturen erfolgreich zu registrieren.

als auch bei multimodalen Aufnahmen gute Registrierungsergebnisse erzielt werden. Somit eignet sie sich gleichermaßen für die Registrierung von Aufnahmen mit unterschiedlichen Modalitäten und Aufnahmen derselben Modalität zu unterschiedlichen Zeitpunkten (Zeitreihen).

Da aus Effizienzgründen ein lokales Optimierungsverfahren verwendet werden muss, ist eine gute Initialisierung der Transformationsparameter nötig. Für den Schädel kann dies einfach durch die in Abschnitt 4.2.1 gefundene global rigide Transformation des Kopfes erfolgen. Je weiter die Knochen jedoch vom Schädel entfernt sind und je stärker die Positionen der Wirbelsäule voneinander abweichen, desto ungeeigneter ist diese Art der Initialisierung. Daher läuft der Registrierungsprozess sukzessive ab. Zuerst wird der Schädel registriert, dann die unmittelbaren Nachbarn (der Unterkiefer und der Halswirbel C1). Dabei wird die Transformation der Nachbarelemente mit der finalen Transformation des Schädels initialisiert. Die erhaltene finale Transformation des Wirbels C1 wird wiederum als Initialisierung für den nächsten Nachbarn, den Halswirbel C2, verwendet. Auf diese Weise wird ein Element nach dem anderen initialisiert und registriert, bis für alle Elemente je eine rigide Transformation  $\mathcal{T}_i$  vorliegt. Die Registrierungsreihenfolge ist in Abbildung 4.34 dargestellt.

Durch Ausnutzen von Anatomiewissen über den Zusammenhang benachbarter Strukturen ist es möglich, auch initial sehr weit voneinander entfernte Strukturen erfolgreich zu registrieren und somit sehr starke Wirbelsäulenpositionsunterschiede auszugleichen. Dieser große Konvergenzbereich wird im Endeffekt durch die Konkatination vieler *kleiner* Transformationen zu einer *großen* Transformation erreicht.

Durch die Abhängigkeit der Transformationen benachbarter Strukturen entsteht jedoch auch ein entscheidender Nachteil. Wurde ein Element nicht korrekt registriert, bspw. aufgrund eines starken Bildartefakts in einer der Aufnahmen an der Struktur, werden alle nachfolgenden Strukturen falsch initialisiert und eine korrekte Registrierung dieser Strukturen wird unwahrscheinlich. Das Ausschließen

von global nicht plausiblen Transformationen, wie dies beim artikulierten Atlas durch den Artikulationsrestriktionsterm erfolgte, könnte hier helfen.

### 4.2.5. Personalisierter artikulierter Atlas

Alternativ zu der lokal rigiden Registrierung soll nun der artikuliert Atlas zunächst personalisiert und anschließend unter Berücksichtigung der möglichen Artikulation an das Modellbild angepasst werden, um für jedes Atlaselement eine rigide Transformation zu finden. Grundvoraussetzung der Personalisierung ist die erfolgreiche Anpassung an eine CT-Aufnahme  $\mathcal{C}$ . Neben den Atlasparametern  $\mathbf{p}$  und  $\mathbf{e}$  sind also auch die Segmentierungen  $\mathcal{S}_i$  bekannt. Gegenüber dem zuvor beschriebenen allgemeinen artikulierten Atlas werden vier patientenspezifische Änderungen vorgenommen:

1. **Form:** Das Wahrscheinlichkeitsbild  $\mathcal{W}_i$  eines Elements, welches aus den Trainingsdaten berechnet wurde, wird durch ein Binärbild ersetzt, welches aus der Segmentierung  $\mathcal{S}_i$  des Elements berechnet wird:  $\mathcal{W}_i(\mathbf{x}) = \mathcal{S}_i(T(\mathbf{e}) \circ T(\mathbf{p}_i) \circ \mathbf{x})$ . Es wird nun also eine konkrete Form, nämlich die des segmentierten Elements, und nicht eine probabilistische Form, verwendet.
2. **Textur:** Bei der Textur wird nun ebenfalls das Durchschnittsbild  $\mathcal{A}_i$  durch einen Ausschnitt aus der CT-Aufnahme des Patienten ersetzt:  $\mathcal{A}_i(\mathbf{x}) = \mathcal{C}(T(\mathbf{e}) \circ T(\mathbf{p}_i) \circ \mathbf{x})$ .
3. **Bildenergie:** Da der personalisierte Atlas sowohl an eine CT-Aufnahme desselben Patienten zu einem anderen Zeitpunkt als auch an eine Aufnahme aus einer anderen Modalität (z.B. MRT) angepasst werden soll, kann der mittlere quadratische Fehler aus Gleichung 4.69 nicht verwendet werden. Stattdessen kommt wie in Abschnitt 4.2.4 die Mutual-Information zum Einsatz. Da für die Mutual-Information nach [114] eine Ableitung vorliegt, kann ein gradientenabstiegbasiertes Anpassungsverfahren weiterhin verwendet werden. Besonderheiten werden später in diesem Abschnitt erläutert.
4. **Anpassung:** Die Größe der Strukturen ist bereits bekannt und es wird angenommen, dass diese über die Intra-Subjekt-Aufnahmen hinweg konstant bleibt. Daher darf sich diese während des Anpassungsprozesses für kein Element ändern. Umgesetzt wird dies, indem bei der Anpassungsiteration gegeben durch Gleichung 4.78 alle Komponenten von  $\mathbf{p}$ , welche sich auf die Größe auswirken, nicht berücksichtigt werden. Somit bleiben diese Komponenten konstant und es wird effektiv nur eine rigide Transformation für jedes Element zugelassen.

Durch diese vier Schritte ist also eine Anpassung des artikulierten Atlas auch an andere Bildmodalitäten als CT möglich, wenn der Atlas zuvor an eine CT-Aufnahme desselben Patienten angepasst wurde. Im Vergleich zu der zuvor beschriebenen sukzessiv lokal rigiden Anpassung hat diese Vorgehensweise den großen Vorteil, dass lokale Registrierungsfehler einzelner Elemente sich nicht auf nachfolgende Elemente auswirken, sondern im Gegenteil, die Nachbarelemente diese Fehler verringern können. Dies geschieht während der Atlasanpassung über das Einbringen von globalem Wissen über die Lage der Elemente zueinander durch den Artikulationsrestriktionsterm. Nicht plausible Transformationen werden ausgeschlossen.

Der kritischste Teil der zuvor beschriebenen vier notwendigen Änderungen zur Personalisierung des artikulierten Atlas ist die Berechnung der Mutual-Information, bei welcher einiges zu beachten ist. Der quadratische Fehler zwischen gelernter Textur und CT-Aufnahme wurde durch Gleichung 4.69 für jedes Element unabhängig berechnet und dann durch Gleichung 4.70 für alle Elemente zu einem Gesamtfehler addiert. Bei der Mutual-Information handelt es sich aber um ein statistisches Maß,



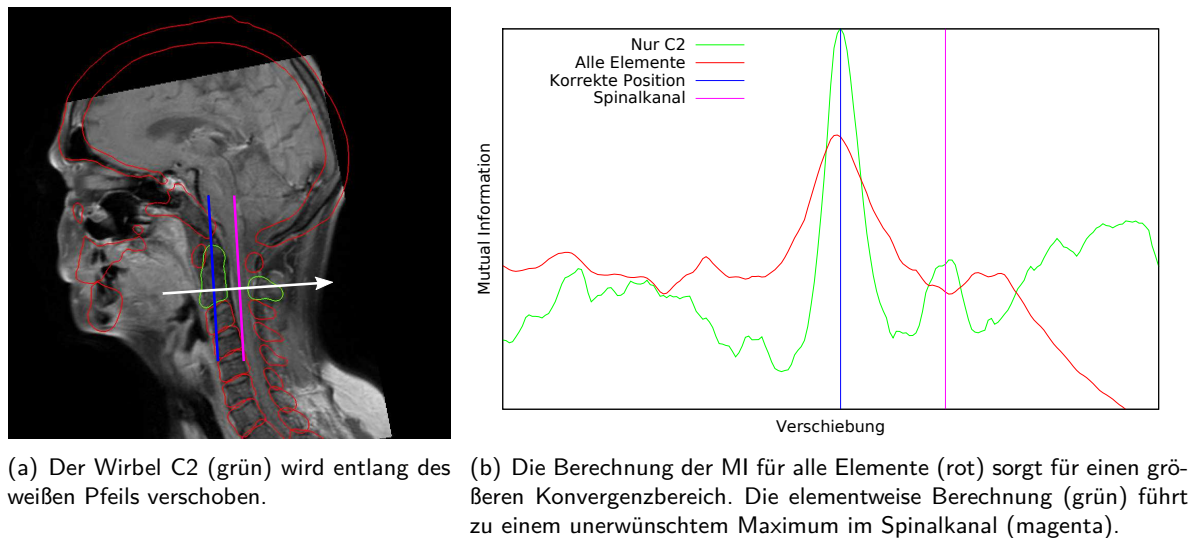


Abbildung 4.35.: Veränderung der Mutual-Information (MI) bei einer Verschiebung des Wirbels C2 bei unterschiedlicher Berechnungsgrundlage

welches lediglich auf beliebige Zusammenhänge zwischen Grauwerten in beiden Aufnahmen setzt und sich somit für multimodale Registrierung eignet. Wird diese allerdings für jedes Element einzeln berechnet, kann es passieren, dass lokal Grauwertzusammenhänge gefunden werden, die global nicht sinnvoll sind. Ein Beispiel dafür ist der Wirbel C2, dessen Zahn parallel zum Spinalkanal verläuft und einen ähnlichen Durchmesser wie dieser besitzt. Wird die Mutual-Information für diesen Wirbel einzeln berechnet, tritt ein lokales Maximum dann auf, wenn sich dessen Zahn im Spinalkanal befindet (siehe Abbildung 4.35). Wird die Mutual-Information hingegen für alle Elemente gleichzeitig berechnet, ist die statistische Grundlage deutlich ausgeprägter. Das Maximum im Spinalkanal verschwindet und der Konvergenzbereich vergrößert sich.

#### 4.2.6. Weichteilgewebe-Deformation

Dieser Abschnitt widmet sich nun der Berechnung des Gesamtdeformationsfeldes aus den  $n$  rigiden Transformationen  $T_i$  der einzelnen Knochen, welche jeweils für die Segmentierungen  $S_i$  im Referenzbild gelten sollen. Dass bei dieser Berechnung keine weitere Bildinformation herangezogen wird, hat folgende Auswirkungen:

- Die Anzahl der Freiheitsgrade ist auf  $6n$  beschränkt. Durch das Einbringen von anatomischem Vorwissen und die globale Regularisierung durch den (personalisierten) artikulierten Atlas werden nur plausible Transformationen erlaubt. Das Verfahren ist somit prinzipbedingt sehr robust gegenüber Bildartefakten.
- Unter der Annahme, dass sich die einzelnen Knochen über die Aufnahmen hinweg nicht anatomisch verändert haben, werden nur Deformationen, die aufgrund einer Änderungen der Lage des Patienten entstanden sind, nicht jedoch diejenigen, welche durch eine Anatomieänderung entstanden sind, ausgeglichen. Dies ist insbesondere bei dem Vergleich einer Zeitserie von Vor-

teil. Rein bildbasierte Verfahren hingegen können zwischen diesen beiden Deformationsquellen nicht unterscheiden.

- Neben diesen Vorteilen hat dieses Vorgehen auch einen Nachteil. Da das Deformationsfeld aufgrund eines Modells berechnet wird, dieses jedoch nicht alle real entstehenden Deformationen abdecken kann, ist ein Genauigkeitsverlust im Weichteilgewebe zu erwarten, welcher vermutlich mit der Distanz zu den Knochen und der Stärke der Deformation zunimmt.

Bei der Berechnung des Transformationsfeldes  $T(\mathbf{x})$  sollen folgende Bedingungen erfüllt sein:

1. Das Transformationsfeld soll kontinuierlich sein. Das heißt, dass  $T(\mathbf{x} + \Delta) \approx T(\mathbf{x})$  für kleine Abweichungen  $\Delta$  gelten soll. Dies bezieht sich allerdings nicht nur auf den Translations- sondern auch auf den Rotationsanteil der Transformation (vgl. Abschnitt 3.2.4).
2. Alle bekannten Transformationen der rigiden Bereiche sollen einbezogen werden und innerhalb der Segmentierungen  $S_i$  dieser Bereiche sollen die bereits bekannten Transformationen  $T_i$  zur Anwendung kommen. Es soll also  $T(\mathbf{x}) = T_i \iff S_i(\mathbf{x}) = 1$  gelten.

Zur Erfüllung dieser beiden Anforderungen wird die Transformation  $T(\mathbf{x})$  am Ort  $\mathbf{x}$  als Linearkombination der bekannten Transformationen  $T_i$  beschrieben, welche daher von nun an als Basistransformationen bezeichnet werden:

$$T(\mathbf{x}) = \bigoplus_{i=1}^n \phi_i(\mathbf{x}) \odot T_i . \quad (4.82)$$

Um Transformationen linear kombinieren zu können, wird, wie schon zur Beschreibung der Transformationen der Elemente des artikulierten Atlas in Abschnitt 4.2.2, auf die Matrixexponentialfunktion zurückgegriffen. Das Koeffizientenfeld  $\phi : \Omega \mapsto [0, 1]^n$  beschreibt, welcher Anteil der jeweiligen Basistransformation  $T_i$  an einem bestimmten Ort enthalten ist. Der Definitionsbereich  $\Omega$  des Koeffizientenfeldes ist dabei nur außerhalb der rigiden Elemente definiert:  $\Omega \subseteq \mathcal{D} \setminus \bigcup_{\forall i}$ .

**Kontinuität durch die Laplacegleichung:** Um Kontinuität des Deformationsfeldes zu erreichen, sollen alle Komponenten von  $\phi$  die Laplacegleichung, eine partielle Differentialgleichung zweiter Ordnung

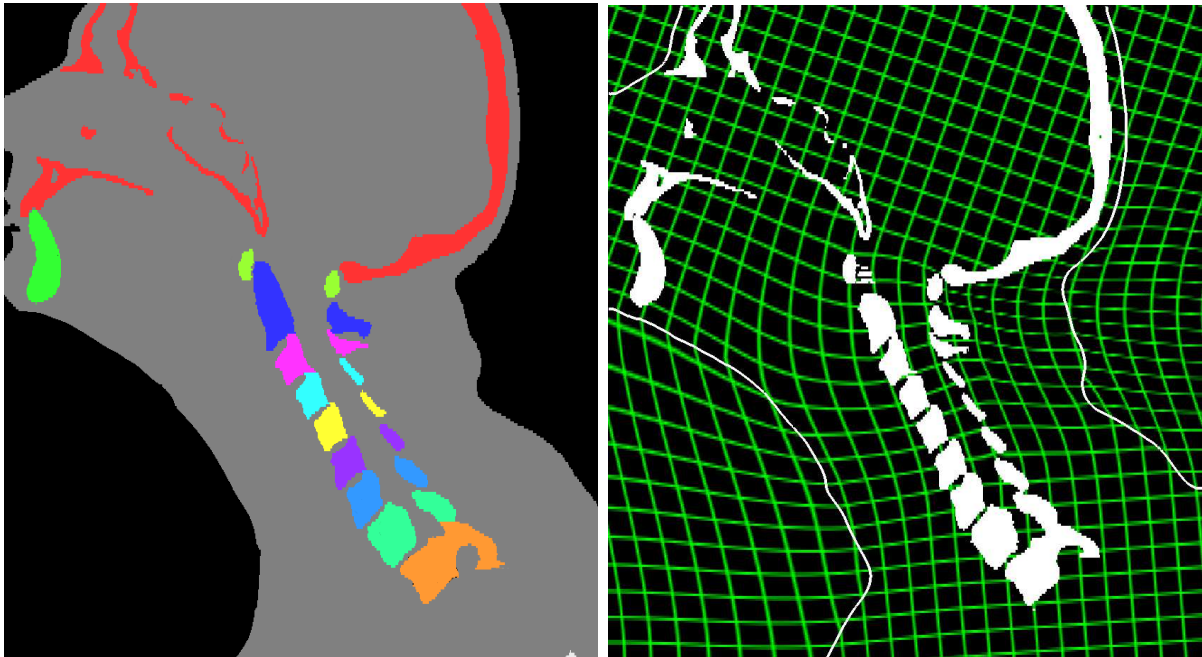
$$\Delta \phi_i = 0 \quad (4.83)$$

erfüllen.  $\Delta = \sum_{\forall j} \frac{\partial^2}{\partial x_j^2}$  ist dabei der Laplaceoperator. Diese partielle Differentialgleichung ist aus der Thermodynamik bekannt, wo sie eingesetzt wird, um eine kontinuierliche Wärmeverteilung zu beschreiben.

**Randbedingungen:** Die Basistransformationen fließen in Form von Dirichlet-Randbedingungen in die Differentialgleichung ein. Am Rand der Segmentierung  $S_i$  soll die  $i$ -te Komponente von  $\phi$  den maximalen Wert von 1 aufweisen und alle anderen Komponenten den minimalen Wert von 0:

$$\phi_j(\mathbf{x}) = \begin{cases} 1 & ; i = j \\ 0 & ; i \neq j \end{cases} \Big| \mathbf{x} \in \partial S_i . \quad (4.84)$$

Die Auswertung von Gleichung 4.82 am Rand von  $S_i$  ergibt also immer  $T(\mathbf{x}) = T_i, \mathbf{x} \in \partial S_i$ . Um eine eindeutige Lösung erhalten zu können, sind weitere Randbedingungen am Rand des Definitionsbereichs der partiellen Differentialgleichung  $\partial\Omega$  nötig. Da dort allerdings die Transformationen nicht



(a) Randbedingungen kommen beim Übergang in den grauen Bereich zum Einsatz (b) Visualisierung des berechneten Deformationsfeldes

Abbildung 4.36.: Das Deformationsfeld wird in zwei Schritten berechnet. Zuerst beschränkt sich der Definitionsbereich auf den Körper (grau) und anschließend erst auf den Rest des Bildes (schwarz).

bekannt sind, können Dirichlet-Randbedingung hier nicht verwendet werden. Stattdessen kommen Neumann-Randbedingungen zum Einsatz, welche die erste Ableitung vorgeben:

$$\frac{\partial \phi_i(\mathbf{x})}{\partial \mathbf{x}} = 0 \quad \left| \quad \mathbf{x} \in \partial \Omega \right. . \quad (4.85)$$

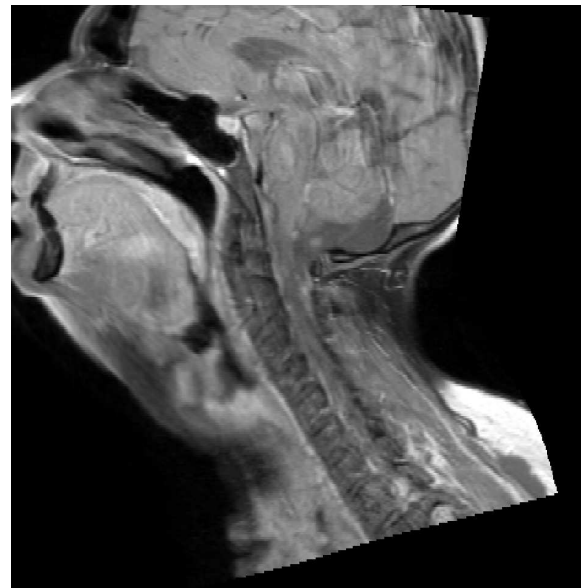
**Lösung:** Die Bestimmung des Transformationsfeldes durch Lösung der Differentialgleichung erfolgt in zwei Schritten.

Im ersten Schritt wird der Definitionsbereich  $\Omega$  auf den Bereich innerhalb des Körpers beschränkt. Dieser wird durch ein Schwellwertverfahren aus der CT-Aufnahme extrahiert (siehe Abbildung 4.36). Diese Beschränkung folgt der Anatomie. Die Transformationen der Knochen werden in das Weichteilgewebe propagiert, nicht jedoch über die Körpergrenze hinaus. Das hat den Vorteil, dass beispielsweise Transformationen des Kinns durch die *Luft* hindurch keinen Einfluss auf die Transformationen im Brustbereich ausüben und so ungewünschte Deformationen vermieden werden (siehe Abbildung 4.37).

Im zweiten Schritt wird der Definitionsbereich auf den Bereich außerhalb des Körpers festgelegt, so dass ein dichtes Gesamtdeformationsfeld berechnet werden kann. Am Körpertrand kommen dann



(a) Die Berechnung des Deformationsfeldes in einem Schritt führt zu ungewünschten Deformationen, insbesondere im Brustbereich (eingekreist). Dies ist dem Einfluss der Transformation des Kinns auf den Brustbereich durch die *Luft* hindurch geschuldet (Pfeil).



(b) Durch die 2-Schritt Anpassung wird das Weichteilgewebe im Hals nicht durch den Kopf beeinträchtigt.

Abbildung 4.37.: Vergleich zwischen unterschiedlichen Lösungsstrategien zur Bestimmung des Deformationsfeldes

Dirichlet-Randbedingungen mit den im vorangegangenen Schritt berechneten Transformationen zum Einsatz. Am Bildrand werden Neumann-Randbedingungen verwendet.

Die eigentliche Lösung der partiellen Differentialgleichung erfolgt dann durch die Finite-Differenzen-Methode. Der Definitionsbereich wird in ein diskretes Voxelgitter aufgeteilt. Jedem der Voxel wird eine Gleichung zugeordnet, welche im Randbereich die Randbedingungen aus Gleichung 4.84 bzw. Gleichung 4.85 und im inneren Bereich der Gleichung 4.83 entspricht. Die Ableitungen werden durch finite Differenzen mit Bezug auf die Nachbarvoxel approximiert. Das resultierende dünn besetzte lineare Gleichungssystem kann dann durch ein Standardverfahren gelöst werden.

Das Ergebnis ist nun eine Transformation  $T(\mathbf{x})$  an jedem Punkt  $\mathbf{x}$ . Unter Anwendung dieser Transformationen kann so das transformierte Modellbild  $M'$  erzeugt werden, welches nun im örtlichen Zusammenhang mit der CT-Aufnahme steht:

$$M'(\mathbf{x}) = M(T(\mathbf{x}) \circ \mathbf{x}), \mathbf{x} \in \mathcal{D}. \quad (4.86)$$

### 4.2.7. Weitere Stützpunkte

Bisher wurde das gesamte Deformationsfeld nur aufgrund der rigid Transformationen der einzelnen Knochen berechnet. Die Vorteile sind eine hohe Robustheit und die Beibehaltung von Anatomieun-

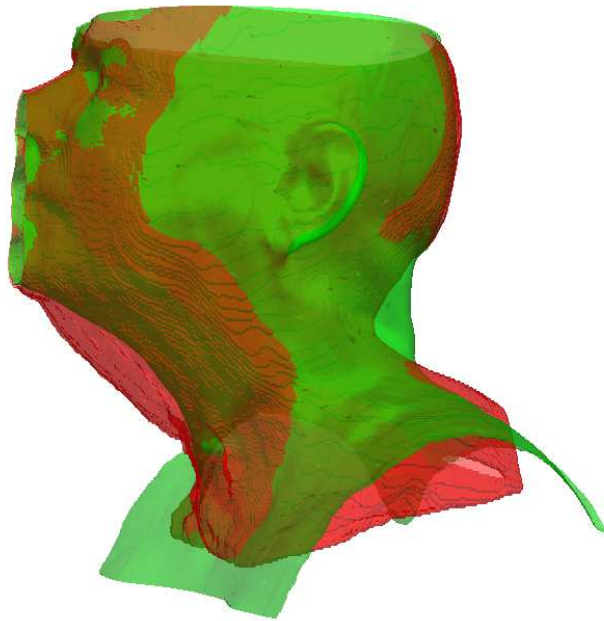
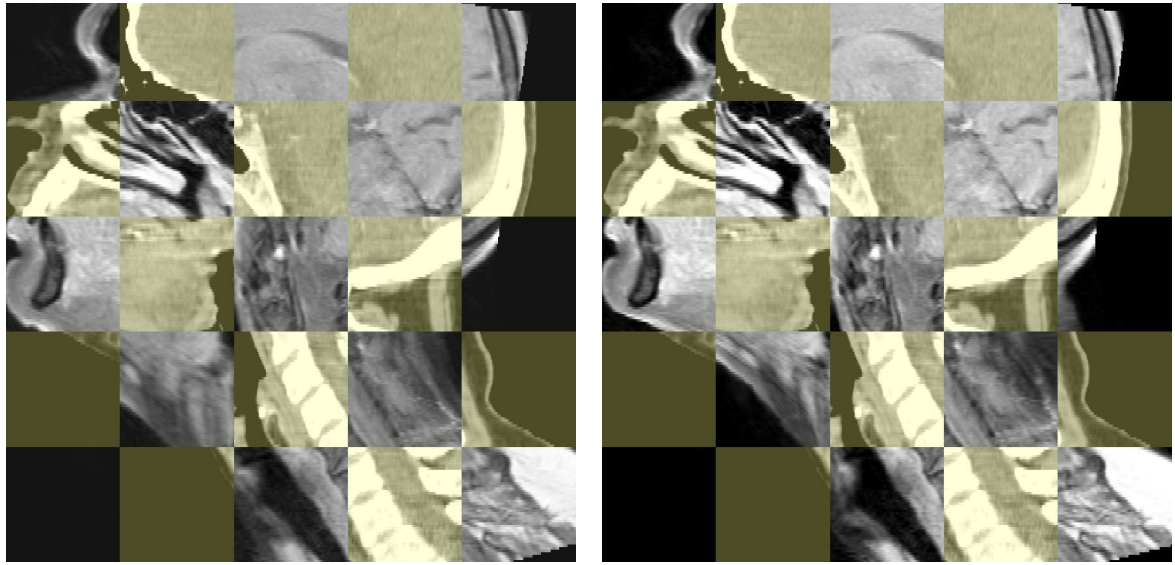


Abbildung 4.38.: Körperoberfläche aus der CT-Aufnahme (grün) und transformierten MRT-Aufnahme (rot)

terschieden in Zeitserienaufnahmen. Der Nachteil hingegen ist, dass die Genauigkeit mit der Distanz zu den Knochen und steigender Deformationsstärke abnimmt. Dieser Nachteil kann durch drei unterschiedliche Ansätze vermindert werden:

1. **Biomechanische Modellierung:** Die einzige Annahme, die bisher zur Berechnung des Transformationsfeldes eingesetzt wurde, ist Kontinuität. Eine biomechanische Modellierung, bspw. durch finite Elemente, könnte möglicherweise die Deformationen realistischer abbilden und so eventuell die Registrierungsgenauigkeit erhöhen. Dazu muss allerdings gewebespezifisches Wissen vorhanden sein, was unter anderem die Segmentierung weiterer Strukturen nach sich zieht.
2. **Bildbasierte Registrierung:** Bei der Berechnung des Transformationfeldes könnte ein Bildähnlichkeitsmaß auch im Weichteilgebe maximiert werden, um auch dort bildbasierte Information bei der Registrierung einfließen zu lassen. Eine mögliche Erhöhung der Genauigkeit würde damit mit den bereits beschriebenen Nachteilen bildbasierter Registrierung – Anfälligkeit gegenüber Bildartefakten, keine Unterscheidung zwischen Deformation aufgrund von Positionsänderung oder Anatomieänderung – einhergehen.
3. **Weitere Stützpunkte:** Zusätzlich zu den Knochen können weitere Strukturen, welche in beiden Aufnahmen robust segmentiert werden können, als weitere Stützpunkte in die Transformationsfeldberechnung einbezogen werden und so den Bereich der zu interpolierenden Region verringern und damit die Genauigkeit erhöhen.

In dieser Arbeit wurde die letzte der drei Möglichkeiten gewählt. Als zusätzliche Struktur wird die Körperoberfläche verwendet. Zum einen ist diese sehr weit vom Skelett entfernt und damit sind dort nach der bisherigen Registrierung die größten Deformationsunterschiede zu erwarten. Zum anderen



(a) Ohne Körperoberflächenanpassung kommt es zu Ungenauigkeiten an Orten, die vom Skelett weit entfernt sind.

(b) Die Körperoberflächenanpassung erhöht die Registrierungsgenauigkeit im Hals- und Nackenbereich.

Abbildung 4.39.: Vergleich des Registrierungsergebnisses mit und ohne zusätzlicher Körperoberflächenanpassung bei einem initial stark deformierten Aufnahmepaar

kann die Körperoberfläche sowohl in CT als auch MRT-Aufnahmen sehr einfach segmentiert werden. In dem Referenzbild geschah dies bereits zur Bestimmung des Definitionsbereichs  $\Omega$  der partiellen Differentialgleichung mit Hilfe eines Schwellwertverfahrens und kann damit wieder verwendet werden. In dem transformierten Modellbild sollte sich die Körperoberfläche bereits in der Nähe der Oberfläche der Referenzaufnahme befinden (siehe Abbildung 4.38). Daher wird von dieser ausgegangen und in Oberflächennormalenrichtung nach starken Bildgradienten in dieser Richtung gesucht. Der Ort mit dem stärksten richtungsgewichteten Bildgradienten wird als Oberflächenpunkt angesehen und es können für diese Orte Deformationsvektoren  $\mathbf{d}(\mathbf{x})$ ,  $\mathbf{x} \in \partial\Omega$  angegeben werden. In den rigiden Elementen wird die Deformation auf 0 gesetzt  $\mathbf{d}(\mathbf{x}) = \mathbf{0}$ ,  $\forall \mathbf{x} | \exists i | \mathcal{S}_i(\mathbf{x}) = 1$ . An allen anderen Orten wird die Deformation  $\mathbf{d}(\mathbf{x})$ , wie auch schon zur Bestimmung des Transformationsfeldes, durch die Lösung der Laplacegleichung bestimmt. Dirichlet-Randbedingungen sorgen dafür, dass die Deformationen an der Körperoberfläche und den Knochen wie beschrieben sind. Am Bildrand kommen Neumann-Randbedingungen zum Einsatz. Nachdem die Lösung wiederum durch finite Differenzen auf einem diskreten Voxelgitter bestimmt wurde, wird Gleichung 4.86 zur Bestimmung des transformierten Modellbildes erweitert:

$$M'(\mathbf{x}) = M(T(\mathbf{x}) \circ \mathbf{x} + \mathbf{d}(\mathbf{x})) . \quad (4.87)$$

Der Unterschied zwischen den Registrierungsergebnissen mit und ohne Einbeziehung der Körperoberfläche ist exemplarisch für eine stark deformierte Aufnahme in Abbildung 4.39 dargestellt. In diesem Fall kann das Registrierungsergebnis durch die zusätzliche Körperoberflächenanpassung insbesondere im stark deformiertem Hals- und Nackenbereich verbessert werden.

Ein Nachteil des Einbringens weiterer Stützpunkte ist allerdings eine geringere Robustheit. Ist bspw. die Oberflächensegmentierung durch das Vorhandensein von starken Bildartefakten (z.B. Metallarte-



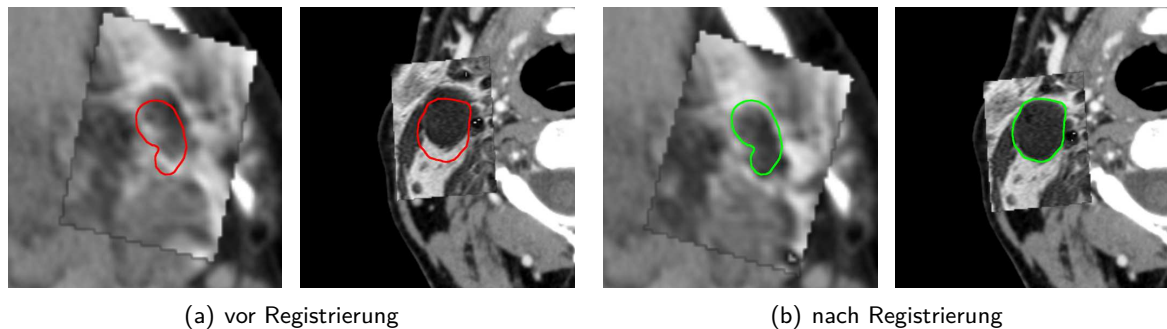


Abbildung 4.40.: Beispiele von Lymphknoten vor und nach lokal rigider Registrierung von CT/MRT-Aufnahmen

fakte in der CT-Aufnahme oder lokal geringe Signalstärke in der MRT-Aufnahme) beeinträchtigt, entstehen nicht gewünschte Deformationen. Fortgeschrittenere Segmentierungsmethoden können diesen Nachteil zum Teil ausgleichen. Eine weitere Eigenschaft, die durch das Einbeziehen der Körperoberfläche verloren geht, ist die Erhaltung von Anatomieunterschieden, weil die Körperoberfläche von Unterschieden im Weichteilgewebe betroffen sein kann. Bspw. kann bei einem Tumorwachstum die Hautoberfläche ausgebeult werden. Wird die Hautoberfläche in den Registrierungsprozess einbezogen, wird diese Ausbeulung im Gegensatz zu dem rein multi-rigiden Verfahren ausgeglichen. Auf diese Eigenschaft sollte bei der Anwendung geachtet werden.

#### 4.2.8. Lokal rigide Registrierung von Lymphknoten

Bisher wurde stets der gesamte Kopf-Hals-Bereich bei der Registrierung betrachtet. Mit Blick auf das Anwendungsszenario dieser Arbeit – Merkmalsextraktion von Lymphknoten aus multimodalen Aufnahmen – soll die Genauigkeit insbesondere innerhalb der Lymphknoten besonders hoch sein. Bisher ist allerdings die Registrierungsgenauigkeit prinzipbedingt nur innerhalb der Knochen besonders hoch. Die Erhöhung der Registrierungsgenauigkeit soll nun durch eine lokal-rigide Registrierungsmethode erreicht werden, die in [16] veröffentlicht wurde. Sie basiert auf zwei Grundannahmen:

1. Lymphknoten sind deutlich härter als das umliegende Weichteilgewebe. Dadurch sind keine Deformationen, sondern lediglich eine Verschiebung und Drehung zu erwarten.
2. Die gesamten Aufnahmen sind bereits global gut vorregistriert, so dass nun noch kleine Verschiebungen und Drehungen zu erwarten sind. Zur Erfüllung dieser Annahme trägt die beschriebene multi-rigide Registrierungsmethode bei.

Eingabe dieser lokal rigiden Registrierung ist eine Segmentierung des zu registrierenden Lymphknotens in einer der Aufnahmen. Aus einer CT-Aufnahme kann diese bspw. durch das in Abschnitt 4.1.3 vorgestellte semi-automatische Lymphknotensegmentierungsverfahren mit hoher Genauigkeit bestimmt werden. Die Segmentierung wird zunächst geringfügig dilatiert und dient dann als Definitionsbereich für eine bildbasierte rigide Registrierung. Wird die Menge an Dilatation zu klein gewählt, sinkt zum einen der Konvergenzbereich und zum anderen ist zu wenig Kontext für eine erfolgreiche Registrierung vorhanden. Wird die Menge an Dilation hingegen zu groß gewählt, steigt die Wahrscheinlichkeit, dass deformierte Bereiche enthalten sind, und somit die Registrierungsgenauigkeit innerhalb des

Lymphknoten sinkt. Eine Dilation von 5 mm in alle drei Raumrichtungen hat sich als geeigneter Kompromiss erwiesen. Die eigentliche Registrierung basiert dann auf der Maximierung der Mutual-Information dieses Bereichs durch Veränderung von rigiden Transformationsparametern durch ein Gradientenabstiegsverfahren. Diese Vorgehensweise ist identisch zu der lokal rigiden Registrierung eines Elements der sukzessiven Registrierung der Knochen aus Abschnitt 4.2.4. In Abbildung 4.40 sind Beispiele von Lymphknoten vor und nach Anwendung des lokal rigide Registrierungsverfahrens dargestellt.

Das beschriebene Verfahren kann sowohl zur multimodalen Registrierung, als auch zur Registrierung von Lymphknoten innerhalb von Zeitserien eingesetzt werden. Im Vergleich zu einer deformierbaren Registrierung ergeben sich durch die lokal rigide Registrierung der Lymphknoten folgende Vorteile:

- Aufgrund der geringeren Anzahl an Freiheitsgraden (3 Translation + 3 Rotation) ist die Komplexität und damit die Laufzeit deutlich geringer.
- Der Lymphknoten wird nicht deformiert. Dadurch ist bei multimodalen Aufnahmen stets sichergestellt, dass die richtigen Strukturen aneinander ausgerichtet werden. Andernfalls könnte bspw. eine zentrale Nekrose, welche in nur einer Aufnahme gut zu erkennen ist, am gesamten Lymphknoten ausgerichtet werden.
- Bei Zeitserien bleiben Anatomieunterschiede erhalten. Wächst der Lymphknoten über die Aufnahmen hinweg, kann dies sehr einfach festgestellt werden, da in so einem Fall keine perfekte Ausrichtung möglich ist.

### 4.2.9. Zusammenfassung

**Methoden:** Zur Registrierung des Kopf-Hals-Bereichs wurde eine vollautomatische global regulierte multi-rigide Methode entwickelt. Nach einer rigiden Registrierung des Kopfes wird ein artikulierter Atlas zur Segmentierung der einzelnen Knochen aus der CT-Aufnahme verwendet. Dieser Atlas basiert auf Wissen über Form und Textur einzelner Knochen und deren örtlichen Zusammenhang (Artikulation), welches aus annotierten Trainings-CT-Aufnahmen gelernt wurde. Eine sukzessiv lokal rigide Registrierung bestimmt einzelne rigide Transformationen der Knochen zu der anderen Aufnahme (andere Modalität oder Zeitserie). Alternativ geschieht dies durch einen personalisierten artikulierten Atlas. Unter der Annahme von Kontinuität wird dann aus diesen Transformationen ein dichtes Gesamttransformationsfeld durch ein zweistufiges Verfahren berechnet. Optional wird die aus beiden Aufnahmen extrahierte Körperoberfläche als weiterer Stützpunkt in den Registrierungsprozess einbezogen. Um eine besonders hohe Registrierungsgenauigkeit in den Lymphknoten zu erreichen, wird dort abschließend ein lokal rigides Registrierungsverfahren angewendet.

**Eigene Beiträge:** Im Vergleich zu existierenden Verfahren ergeben sich folgende Neuerungen:

- **Vollautomatisches Verfahren:** Im Gegensatz zu allen existierenden multi-rigiden Kopf-Hals-Registrierungsverfahren ist das im Rahmen dieser Arbeit entwickelte Verfahren vollautomatisch. In der Literatur stellt dabei die automatische Segmentierung der einzelnen rigiden Teile die größte Herausforderung dar, was in dieser Arbeit durch den neuartigen artikulierten Atlas geschieht.
- **Robuste rigide Registrierung:** Im Gegensatz zum Standardverfahren zur multimodalen rigi-



den Registrierung ist das in dieser Arbeit vorgestellte Verfahren durch die automatische Extraktion einer gemeinsamen Landmarke in beiden Aufnahmen in der Lage, sowohl mit starken Deformationen des Halses als auch mit stark beschnittenen Aufnahmen zurechtzukommen.

- **Gelernte Artikulation:** Der in dieser Arbeit vorgestellte artikulierte Atlas lernt die örtlichen Zusammenhänge mehrerer Strukturen aus Trainingsdaten. In der Literatur geschieht dies, wenn überhaupt, üblicherweise durch eine explizite Modellierung.
- **Adaptive Anpassung:** Bei der Anpassung des artikulierten Atlas ändert sich der Einfluss der Bildenergie für jedes Element im Laufe der Anpassung, wodurch eine lokale Anpassung an falsche Strukturen vermieden wird. In der Literatur erfolgt die Anpassung üblicherweise für alle Strukturen gleichzeitig.
- **Personalisierter artikulierter Atlas:** Die Personalisierung des artikulierten Atlas ermöglicht eine globale Regularisierung zur Bildregistrierung. Darüber hinaus ermöglicht es auch die erfolgreiche Segmentierung von sehr schwach kontrastierten Strukturen, wie z.B. Knochen in MRT-Aufnahmen.
- **Deformationsfeldberechnung:** Im Gegensatz zu den meisten interpolationsbasierten Deformationsfeldberechnungsmethoden ist die Linearkombination von Transformationen durch die Verwendung des Matrixexponentials mathematisch fundiert. Während des zweistufigen Prozesses wird bedingt Wissen über die Deformation von unterschiedlichem Material (Gewebe/Luft) eingebracht.

**Vorteile:** Gegenüber rigiden Registrierungsverfahren hat das in dieser Arbeit vorgestellte Verfahren den Vorteil, dass auch Aufnahmen des Kopf-Hals-Bereichs erfolgreich registriert werden können, in denen die Wirbelsäule in unterschiedlichen Positionen ist, es also zu Deformationen im Weichteilgewebe kommt.

Gegenüber rein bildbasierten nichtrigiden Registrierungsverfahren, welche üblicherweise eingesetzt werden, um solche Deformationen auszugleichen, hat es zwei Hauptvorteile: ein großer Konvergenzbereich und eine hohe Robustheit.

Der große Konvergenzbereich wird durch die Verwendung von Wissen über die anatomische Lage der Knochen zueinander erreicht. Auf dieses Wissen wird bei der Segmentierung der Knochen durch den artikulierten Atlas und dessen adaptiver Anpassung, aber auch bei der sukzessiven lokalen rigiden Registrierung der Knochen zurückgegriffen. Das Lernen des Artikulationsraums aus Trainingsdaten hat den Vorteil, dass das Verfahren trotz der hohen Komplexität vergleichsweise generisch ist. Die Anwendung auf eine andere Körperregion erfordert lediglich ein Training mit entsprechenden Trainingsdaten. Die Entwicklung eines neuen Verfahrens ist nicht notwendig.

Die hohe Robustheit wird durch die – für ein deformierbares Verfahren – geringe Anzahl an Transformationsparametern, die Verwendung von globaler Regularisierung und der nicht bildbasierten Bestimmung des Deformationsfeldes erreicht. Das Vorhandensein von Bildartefakten führt daher nicht zu unnatürlichen Deformationen.

**Nachteile und Lösungsansätze:** Die Möglichkeit, Aufnahmen in unterschiedlicher Wirbelsäulenstellung mit hoher Genauigkeit registrieren zu können, wird durch eine deutlich höhere Komplexität im Vergleich zu rigider Registrierung erkauft. Viele, teils komplexe Schritte sind nötig. Schlägt ein

Schritt fehlt, können auch die folgenden Schritte nicht korrekt ausgeführt werden und die gesamte Registrierung schlägt fehl.

Auch die globale Regularisierung wirkt sich in Teilbereichen nachteilig aus. Ist eine auftretende Deformation nicht ausreichend durch die Trainingsdaten repräsentiert, kann kein zufriedenstellendes Registrierungsergebnis erzielt werden. Dies kann insbesondere bei Pathologien der Fall sein. Eine repräsentativere Trainingsdatenbasis und das Tolerieren einer gewissen Abweichung vom gelernten Transformationsraum kann dem entgegenwirken.

Darüber hinaus ist das in dieser Arbeit vorgestellte Verfahren im Vergleich zu rein bildbasierten Verfahren nicht direkt auf beliebige Körperregionen anwendbar, sondern es bedarf des Lernens des spezifischen Transformationsraums aus annotierten Trainingsdaten. Ein weiterer Nachteil besteht in der möglicherweise geringeren Registrierungsgenauigkeit im Weichteilgewebe, weil dessen Bildintensitäten nicht während des Registrierungsprozesses berücksichtigt werden. Auf der anderen Seite wird dadurch aber eine Zeitserienanalyse ermöglicht, weil Unterschiede in der Anatomie des Weichteilgewebes nicht ausgeglichen werden, was gewollt war. Das Nachschalten eines lokal rigiden Verfahrens für die Zielstrukturen (hier Lymphknoten) kann einen etwaigen Genauigkeitsverlust zumindest für diese Strukturen prinzipiell ausgleichen.

## 5. Experimente & Ergebnisse

### 5.1. Segmentierung von Lymphknoten

#### 5.1.1. Material

Grundlage für die Evaluation bildet eine Menge bestehend aus 49 Lymphknoten, für die sowohl eine CT-Aufnahme als auch eine Referenzsegmentierung vorliegt.

**Auswahl der Daten:** Eine erfahrene Radiologin (Inmaculada Rodriguez Jimenez aus der Abteilung für Radiologie des Krebszentrums MD Anderson in Madrid, Spanien mit 15-jähriger Berufserfahrung) wählte insgesamt 49 Lymphknoten aus 13 kontrastverstärkten CT-Aufnahmen des Kopf-Hals-Bereichs von Patienten mit Kopf-Hals-Karzinomen aus. Unter den Patienten befanden sich 5 Männer und 8 Frauen im Alter von 36 bis 93 Jahren. Die CT-Aufnahmen wurden im Universitätsklinikum in Parma im klinischen Alltag mit dort üblichen Einstellungen erstellt. Der Schichtabstand variiert zwischen 0.5 und 1.2 *mm*.

In der Menge befinden sich sowohl gutartige als auch bösartige Lymphknoten. Letztere sind teilweise stark vergrößert und besitzen nekrotisches Gewebe im Inneren. Die Form reicht von der typischen Bohnenform gesunder Lymphknoten bis hin zu der sehr runden Form vergrößerter Lymphknoten. Das umliegende Gewebe ist sehr unterschiedlich und reicht von Knochen über kontrastverstärkte Gefäße, Drüsen, Muskeln, weitere Lymphknoten und Fett bis hin zu Luft. Folglich ist der Kontrast zum umliegenden Gewebe örtlich sehr gering. Zusätzlich weist ein Teil der CT-Aufnahmen erhebliches Rauschen auf.

Tabelle 5.1 enthält einige statistische Aussagen über die vorliegende Menge. In den Abbildungen aus Kapitel 2.1 sind ausgewählte, auf unterschiedliche Weise interessante Lymphknoten zu sehen. Laut Aussage der klinischen Partner kann die vorliegende Menge als repräsentativ für Lymphknoten des Kopf-Hals-Bereichs betrachtet werden. Alle für die Lymphknoten-Segmentierung schwierigen Fälle werden abgedeckt.

**Erstellung der Referenzsegmentierung:** Die Messung der Segmentierungsgenauigkeit stellt in der medizinischen Bildverarbeitung eine große Herausforderung dar, weil in der Regel keine Referenzsegmentierung vorliegt, gegen die ein Segmentierungsergebnis eines Algorithmus gemessen werden kann. Lediglich bei künstlich generierten Daten entspricht die Referenzsegmentierung auch der tatsächlichen Wahrheit, hat aber den entscheidenden Nachteil, dass keine belastbare Aussage über die Segmentierungsgenauigkeit auf klinischen Daten getroffen werden kann.

Notgedrungen hat es sich daher durchgesetzt, eine durch einen Experten manuell erstellte Segmentierung als Referenzsegmentierung anzusehen. Diese unterliegt jedoch einer Subjektivität. Nicht nur

|                                    | Min  | Max   | Ø                 |
|------------------------------------|------|-------|-------------------|
| Volumen (ml)                       | 0.22 | 27.38 | $1.70 \pm 3.92$   |
| großer Durchmesser (mm)            | 9.18 | 47.55 | $17.11 \pm 6.95$  |
| kleiner Durchmesser (mm)           | 5.21 | 32.51 | $9.96 \pm 4.92$   |
| Verhältnis der Durchmesser         | 1.21 | 2.71  | $1.78 \pm 0.35$   |
| CT-Intensitäten (alle)             | −56  | 240   | $47.55 \pm 35.41$ |
| CT-Intensitäten (ohne nek. Gewebe) | −49  | 194   | $80.90 \pm 19.93$ |

Tabelle 5.1.: Statistische Aussagen über die zur Evaluation verwendete Menge aus 49 Lymphknoten

verschiedene Experten kommen zu verschiedenen Ergebnissen (Inter-Beobachter-Variabilität), sondern auch derselbe Experte bei wiederholter Segmentierung derselben Struktur (Intra-Beobachter-Variabilität).

Besser ist es, einen Konsens aus möglichst vielen manuellen Segmentierungsvorschlägen zu generieren und diesen dann als Referenzsegmentierung zu betrachten. In diesem Fall wurden alle 49 Lymphknoten von 5 verschiedenen Radiologen aus 2 unterschiedlichen Einrichtungen segmentiert. Durchschnittlich benötigte ein Radiologe 4.5 Minuten für die Segmentierung eines einzelnen Lymphknotens, im schlechtesten Fall sogar bis zu 25 Minuten. Aus den 5 Segmentierungsvorschlägen wurde die Referenzsegmentierung dann durch den *Simultaneous Truth and Performance Level Estimation*-Algorithmus [69] (STAPLE) erstellt.

Dieser Algorithmus basiert auf dem Prinzip der *Expectation Maximization*. Jedem Segmentierungsvorschlag werden zunächst unbekannte Parameter über die Qualität der Segmentierung zugeordnet (Sensitivität und Spezifität). Im Expectation-Schritt werden diese Parameter als gegeben angenommen und daraus dann die wahrscheinlichste Referenzsegmentierung berechnet. Im Maximization-Schritt wird diese Referenzsegmentierung als gegeben gesehen und daraus dann die Qualitätsparameter der Segmentierungsvorschläge geschätzt.

**Ähnlichkeits- / Fehlermaße:** Die Segmentierungsgenauigkeit wird durch einen Vergleich des Segmentierungsergebnisses mit einer Referenzsegmentierung bestimmt, welcher durch die Berechnung eines Ähnlichkeits- bzw. Fehlermaßes erfolgt. Im Folgenden werden einige übliche Ähnlichkeits- und Fehlermaße vorgestellt, welche in dieser Arbeit verwendet werden.

Seien  $A$  und  $B$  zwei Mengen, die jeweils alle Voxel einer möglichen Segmentierung enthalten. Weiterhin sei  $d(x, y)$  die euklidische Distanz zwischen den Voxeln  $x$  und  $y$ .

Der *Dice Similarity Coefficient* (DSC) ist ein häufig verwendetes Ähnlichkeitsmaß und ist durch

$$\text{DSC}(A, B) = \frac{2|A \cap B|}{|A| + |B|} \quad (5.1)$$

gegeben. Gilt  $A = B$ , sind Zähler und Nenner gleich und der DSC ergibt sich zu 1. Bei disjunkten Mengen  $A$  und  $B$  hingegen ergibt sich der Zähler und somit auch der DSC zu 0. Der DSC ist also ein volumetrisches Maß für die Ähnlichkeit zweier Segmentierungen, dessen Werte sich im Bereich  $[0, 1]$  befinden.

Ein volumetrisches Fehlermaß ist der *Volumetric Overlap Error* (VOE), welcher den prozentualen

Bestandteil des sich nicht überlappenden Bereichs angibt:

$$\text{VOE}(A, B) = \left(1 - \frac{|A \cap B|}{|A \cup B|}\right) \cdot 100. \quad (5.2)$$

Die Werte des VOE befinden sich im Bereich von  $[0, 100]$ , wobei der Wert 0 nur bei identischen Segmentierungen erreicht wird.

Da beide Maße auf dem Verhältnis von Voxelmengen basieren, werden bei kleineren Objekten mit wenigen Voxeln üblicherweise schlechtere Ergebnisse erreicht als bei großen Objekten. Darüber hinaus wird der Ort des Fehlers nicht einbezogen. Ungenauigkeiten, die sich an der Objektgrenze befinden, wirken sich genauso auf das Maß aus, wie Voxel, die sich in einem komplett anderen Bildbereich befinden.

Häufig verwendete Fehlermaße, welche den Abstand der Elemente der Mengen einbeziehen, sind die symmetrische *Average Surface Distance* (ASD) und die symmetrische *Maximum Surface Distance* (MSD):

$$\text{ASD}(A, B) = \frac{\frac{1}{|A|} \sum_{\forall a \in A} \min_{\forall b \in B} d(a, b) + \frac{1}{|B|} \sum_{\forall b \in B} \min_{\forall a \in A} d(a, b)}{2}, \quad (5.3)$$

$$\text{MSD}(A, B) = \max \left( \max_{\forall a \in A} \min_{\forall b \in B} d(a, b), \max_{\forall b \in B} \min_{\forall a \in A} d(a, b) \right). \quad (5.4)$$

Letztere ist auch unter dem Begriff *Hausdorff-Distanz* bekannt und ist sehr sensitiv gegenüber Ausreißern. Beide Maße befinden sich im Bereich  $[0, \infty[$ , wobei 0 nur für identische Mengen  $A$  und  $B$  angenommen wird.

Da alle Maße unterschiedliche Aspekte der Segmentierungsgenauigkeit betrachten und kein einzelnes Maß für geeigneter als die anderen betrachtet werden kann, sollen alle vier Maße zu einem Index  $\mathcal{G}$  kombiniert werden, welcher die Segmentierungsgüte repräsentiert. Dieser Index ist analog zu dem in [150] vorgestellten Index definiert. Zunächst werden alle Maße so skaliert, dass der Wert 100 perfekter Überlappung und der Wert 75 der Qualität manueller Segmentierung entspricht. Für die ASD bedeutet dies bspw.  $\mathcal{G}_{\text{ASD}} = \max \left( 0, 100 - 25 \cdot \frac{\text{ASD}}{\text{ASD}_{\text{Manuell}}} \right)$ . Das Abschneiden an dem Wert 0 hat den Effekt, dass negative Ausreißer nicht einen unverhältnismäßig starken Einfluss haben. Der Index  $\mathcal{G}$  ergibt sich dann durch Mittelung dieser transformierten Maße zu:

$$\mathcal{G} = \frac{1}{4} (\mathcal{G}_{\text{DSC}} + \mathcal{G}_{\text{VOE}} + \mathcal{G}_{\text{ASD}} + \mathcal{G}_{\text{MSD}}). \quad (5.5)$$

**Inter-Observer-Variabilität:** Die Berechnung des Qualitätsindex  $\mathcal{G}$  setzt voraus, dass die Qualität manueller Segmentierung bekannt ist. Diese wurde durch *Leave-One-Out-Kreuzvalidierung* bestimmt. Dazu wurde eine Referenzsegmentierung aus 4 der Segmentierungsvorschläge durch STAPLE bestimmt und anschließend die oben vorgestellten Qualitätsmaße zum übrigen Segmentierungsvorschlag berechnet. Dieser Vorgang wurde für alle möglichen Kombinationen der Segmentierungsvorschläge wiederholt und die berechneten Qualitätsmaße wurden gemittelt. Die Ergebnisse sind in Tabelle 5.2 eingetragen.

|             | DSC  | VOE   | ASD  | MSD   | $\mathcal{G}$ |
|-------------|------|-------|------|-------|---------------|
| Minimum     | 0.24 | 10.54 | 0.14 | 1.00  | -             |
| Median      | 0.86 | 24.02 | 0.31 | 2.12  | -             |
| Maximum     | 0.94 | 86.11 | 2.75 | 18.18 | -             |
| $\emptyset$ | 0.83 | 27.65 | 0.42 | 3.18  | 75            |

Tabelle 5.2.: Segmentierungsgenauigkeit manueller Segmentierung, welche durch Leave-One-Out-Kreuzvalidierung aus den 5 verschiedenen Expertensegmentierungen bestimmt wurde

**Kreuzvalidierung:** Je größer die Evaluationsmenge ist, desto verlässlicher sind auch statistische Aussagen, die getroffen werden können. Dasselbe gilt für die Menge, die in Kapitel 4.1 zur Parameterbestimmung eingesetzt wurde. Beide Mengen dürfen sich jedoch niemals überschneiden, da sonst eine unabhängige, aussagekräftige Evaluierung nicht möglich ist. Um beiden konträren Anforderungen dennoch gerecht zu werden, wurde in der gesamten Arbeit 3-fache Kreuzvalidierung eingesetzt. Dazu wurde die Gesamtmenge in drei Teilmengen derselben Größe aufgeteilt. Um zu gewährleisten, dass jede dieser Mengen für sich möglichst repräsentativ ist, wurden Lymphknoten mit nekrotischem Gewebe, Lymphknoten mit geringem Bildkontrast, Lymphknoten in stark verrauschten Bildern und alle weiteren Lymphknoten getrennt zufällig auf die Teilmengen verteilt. Diese Teilmengen werden nun als Testmengen  $\mathcal{T}_1$ ,  $\mathcal{T}_2$  und  $\mathcal{T}_3$  bezeichnet. Jeder Testmenge  $\mathcal{T}_i$  wird eine Trainingsmenge  $\mathcal{T}_i^*$  zugeordnet, welche alle Elemente enthält, die nicht in  $\mathcal{T}_i$  selbst enthalten sind. Es gilt also bspw.  $\mathcal{T}_1^* = \mathcal{T}_2 \cup \mathcal{T}_3$ .

In der gesamten Arbeit erfolgte sämtliche Bestimmung von Parametern unabhängig voneinander 3 mal auf den Trainingsmengen. Die Evaluation fand dann ebenfalls dreimal auf den entsprechenden Testmengen statt. Die später in diesem Kapitel berichteten quantitativen Werte über die Segmentierungsgenauigkeit sind aus den drei Mengen kombiniert. Durch diese Vorgehensweise ist sichergestellt, dass alle Lymphknoten sowohl zur Parameterbestimmung als auch zur Evaluation verwendet wurden, es jedoch niemals zu Überschneidung zwischen Trainings- und Testmenge kommt.

### 5.1.2. Größeninvariantes Masse-Feder-Modell

Die Anfälligkeit des Masse-Feder-Modells in Bezug auf die Größe des zu segmentierenden Objektes war die Motivation für die in Abschnitt 4.1.1 beschriebenen Erweiterungen. In diesem Abschnitt wird nun überprüft, inwiefern diese Erweiterungen tatsächlich zur Größeninvarianz beitragen. Diese Evaluation fand auf lediglich 25 Lymphknoten, deren großer Durchmesser sich zwischen 9.2 und 22.3 mm befand, statt, weil die oben beschriebene Menge aus 49 Lymphknoten zum Zeitpunkt der Experimentdurchführung noch nicht verfügbar war.

Diese wurden mit dem entwickelten erweiterten Modell (EMSM) segmentiert und der durchschnittliche DSC wurde berechnet. Zum Vergleich wurde das Masse-Feder-Modell ohne die beschriebenen Erweiterungen (MSM) herangezogen. Unterschiedliche initiale Modellgrößen (4 mm, 10 mm, 15 mm und 20 mm) wurden dabei verwendet. Zusätzlich wurde für jeden Lymphknoten einzeln retrospektiv dasjenige Modell identifiziert, welches die größte Segmentierungsgenauigkeit erzielte. Die Ergebnisse sind in Tabelle 5.3 dargestellt.

Trotz der Verwendung unterschiedlicher Modellgrößen des Vergleichsverfahrens konnten die Erweiterungen in 20 der 25 Fälle bessere Ergebnisse erzielen. Die geringere Standardabweichung verdeutlicht

| Verfahren  | DSC-Durchschnitt | Std. Abweichung |
|------------|------------------|-----------------|
| MSM 4 mm   | 0.43             | 0.16            |
| MSM 10 mm  | 0.58             | 0.16            |
| MSM 15 mm  | 0.42             | 0.18            |
| MSM 20 mm  | 0.23             | 0.12            |
| Bestes MSM | 0.61             | 0.15            |
| EMSM       | 0.72             | 0.11            |

Tabelle 5.3.: Vergleich der Segmentierungsgenauigkeit des erweiterten Masse-Feder-Modells (EMSM) mit dem Standard-MSM verschiedener initialer Größen für 25 Lymphknoten

|             | DSC  | VOE   | ASD   | MSD   | $\mathcal{G}$ |
|-------------|------|-------|-------|-------|---------------|
| Minimum     | 0.11 | 24.83 | 0.31  | 1.22  | 3.18          |
| Median      | 0.72 | 44.21 | 0.84  | 4.82  | 57.19         |
| Maximum     | 0.86 | 99.45 | 11.20 | 24.40 | 79.68         |
| $\emptyset$ | 0.65 | 49.66 | 1.28  | 5.58  | 52.09         |

Tabelle 5.4.: Segmentierungsgenauigkeit des erweiterten Masse-Feder-Modells (EMSM) für alle 49 Lymphknoten

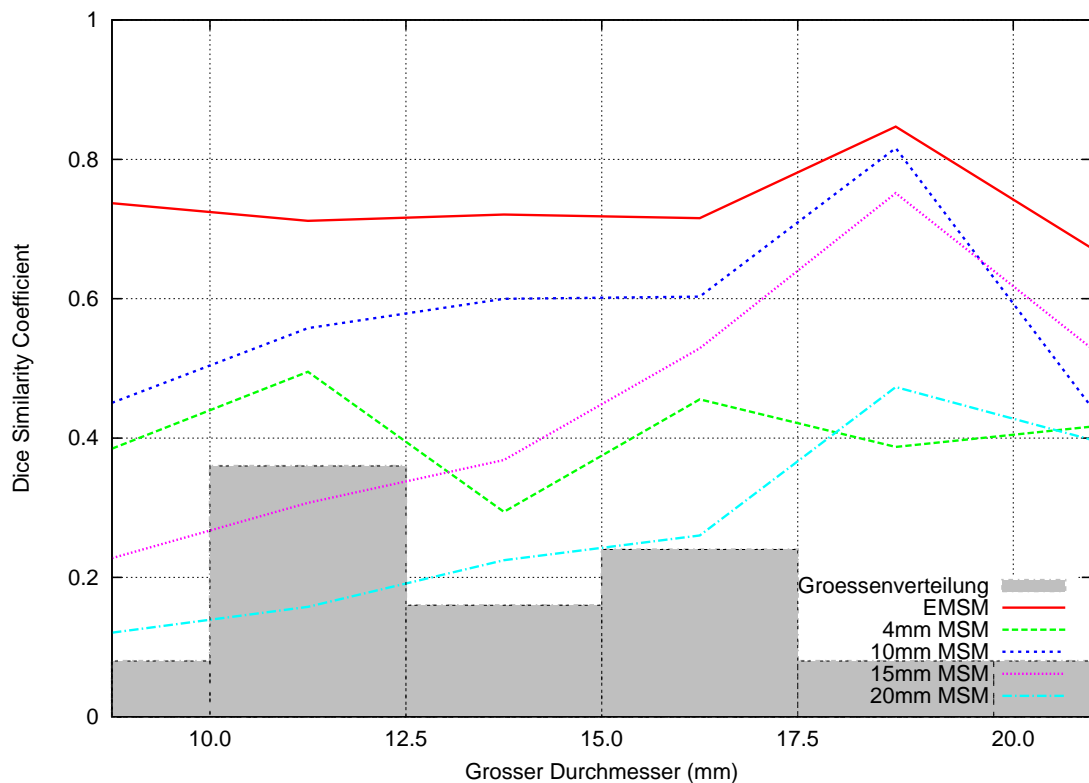


Abbildung 5.1.: Abhängigkeit der Segmentierungsgenauigkeit von der Größe des zu segmentierenden Objektes

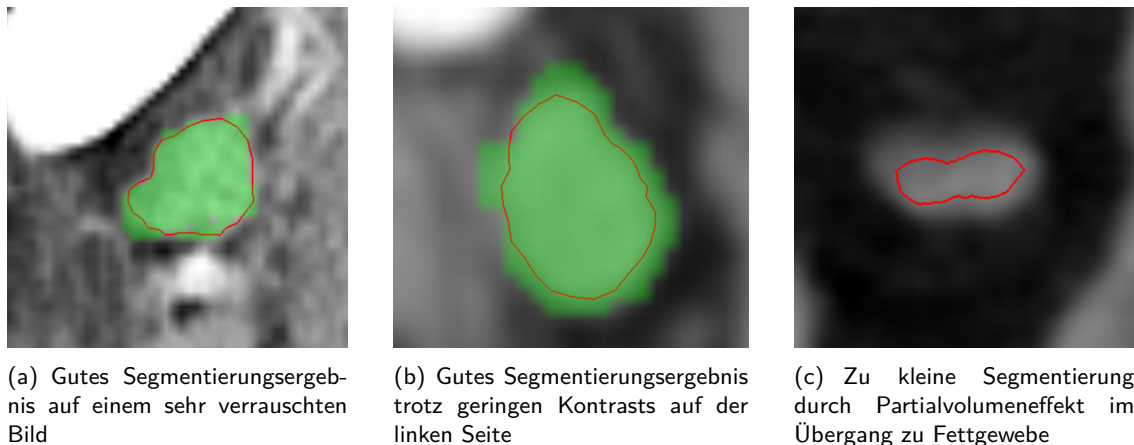


Abbildung 5.2.: Exemplarische Segmentierungsergebnisse des Radialstrahl-basierten 2D-Segmentierungsverfahrens (rot) im Vergleich zur Referenzsegmentierung (grün)

weiterhin die Robustheit gegenüber der Objektgröße. In Abbildung 5.1 ist die Segmentierungsgenauigkeit mit der Größe des Lymphknotens in Bezug gesetzt. Die beschriebenen Erweiterungen führen zu besseren Segmentierungsergebnissen. Darüber hinaus zeigt sich, dass Modelle, welche initial deutlich kleiner als die zu segmentierende Struktur sind, sich besser als größere Modelle eignen. Eine schärfere Aussage kann aufgrund der kleinen Datenbasis nicht getroffen werden.

Später wurde das größeninvariante Feder-Masse-Modell anhand des kompletten im vorangegangenen Abschnitt beschriebenen Datensatzes evaluiert. Die Ergebnisse sind in Tabelle 5.4 zu sehen.

Trotz der sinnvollen Erweiterungen befindet sich die Segmentierungsgenauigkeit nicht annähernd im Bereich der manuellen Segmentierung ( $\mathcal{G} = 75$ ). Ein großes Problem ist die lokale Modellanpassungsstrategie. Wird das Modell von den falschen Bildgradienten angezogen, ist es fast unmöglich, diesen zu entkommen. Auch ist die Verwendung von globalem Formwissen (Kugelform) nicht angebracht, weil die Form von Lymphknoten sehr stark variiert, ohne eine bestimmte globale Form zu bevorzugen.

### 5.1.3. Segmentierung von Lymphknoten in einzelnen CT-Schichten

Evaluationsgrundlage des Radialstrahl-basierten 2D-Segmentierungsverfahrens ist ebenfalls die oben beschriebene Menge bestehend aus 49 Lymphknoten. Das Verfahren wurde jeweils auf eine zentrale Schicht des Lymphknotens angewendet. Die Anzahl der Strahlen wurde dabei auf 32 und die Anzahl der möglichen Radien für jeden Strahl auf 60 festgelegt. Dann wurden Ähnlichkeits- und Fehlermaße im Vergleich zur Referenzsegmentierung berechnet und mit manueller Segmentierung von Experten verglichen. Die erzielten Werte sind in Tabelle 5.5 eingetragen und einige exemplarische Segmentierungsergebnisse sind in Abbildung 5.2 zu sehen.

Die Segmentierungsgenauigkeit ist bei den meisten Maßen nur etwas schlechter als die von manueller Segmentierung. Bei der maximalen ASD ist das Ergebnis sogar leicht besser.

Durch die Verwendung eines globalen Optimierungsverfahrens gelingt es, Lymphknoten unterschiedlicher Größen zu segmentieren. Das lokale Formwissen ermöglicht es, auch bei lokal geringem Bildkon-



|   |             | DSC  | VOE   | ASD  | MSD   | $\mathcal{G}$ |
|---|-------------|------|-------|------|-------|---------------|
| Radialstrahl-basierte<br>2D-Segmentierung | Min         | 0.50 | 11.17 | 0.03 | 1.00  | -             |
|   | Max         | 0.94 | 66.67 | 0.86 | 11.85 | -             |
|   | $\emptyset$ | 0.81 | 31.24 | 0.18 | 2.48  | 56.89         |
| Manuelle Segmentierung                    | Min         | 0.58 | 7.68  | 0.03 | 0.50  | -             |
|   | Max         | 0.96 | 58.80 | 1.58 | 11.51 | -             |
|   | $\emptyset$ | 0.89 | 20.03 | 0.09 | 1.54  | 75.00         |

Tabelle 5.5.: Vergleich der Segmentierungsgenauigkeit des Radialstrahl-basierten 2D-Segmentierungsverfahrens und manueller Segmentierung

trast ein plausibles Ergebnis zu erzielen. Die auf kumulierten Bildintensitäten und dem Bildgradienten basierende Kostenfunktion hat sich bewährt. Es werden sowohl Lymphknoten mit zentralem nekrotischen Gewebe, als auch Lymphknoten, welche durch starkes Rauschen überlagert sind, problemlos segmentiert.

Lediglich ein kleiner Lymphknoten (siehe Abbildung 5.2(c)) zeigt eine systematische Abweichung, obwohl eigentlich ausreichend Bildkontrast vorhanden ist. Eine Analyse dieses Falls ergab, dass die Grauwerte im Übergang zum umliegenden Fettgewebe vermutlich aufgrund des Partialvolumeneffektes geringer sind als die Intensitätskosten erlauben. Eine Vergrößerung des zugelassenen Grauwertbereichs führte allerdings zu einer Übersegmentierung mehrerer anderer Lymphknoten und damit zu einer geringeren durchschnittlichen Segmentierungsgenauigkeit. Abhilfe schaffte die Modellierung eines systematischen Fehlers, ähnlich zu dem in Abschnitt 4.1.5 für das 3D-Verfahren eingeführten Parameter  $\delta$ . Auf dem betrachteten Lymphknoten konnte ein deutlich besseres Ergebnis erzielt werden und sogar der gesamte durchschnittliche DSC stieg auf 0.84.

Zusammenfassend kann gesagt werden, dass ein Radialstrahl-basiertes 2D-Verfahren zur Lymphknotensegmentierung in CT-Schichten geeignet ist. Eine Erweiterung auf den 3D-Fall ist vielversprechend.

#### 5.1.4. Radialstrahl-basierte Segmentierung

**Theorisches Limit:** Prinzipbedingt muss bei der Radialstrahl-basierten Segmentierung jeder Punkt auf der Objektoberfläche auf direktem Weg mit dem Saatpunkt verbunden werden können,

|             | DSC             | VOE             | ASD             | MSD             | $\mathcal{G}$    |
|-------------|-----------------|-----------------|-----------------|-----------------|------------------|
| Minimum     | 0.90            | 0.00            | 0.00            | 0.00            | 79.79            |
| 0.25-Quant. | 1.00            | 0.16            | 0.00            | 0.44            | 97.42            |
| Median      | 1.00            | 0.49            | 0.00            | 0.81            | 98.11            |
| 0.75-Quant. | 1.00            | 0.98            | 0.01            | 1.10            | 99.04            |
| Maximum     | 1.00            | 17.95           | 0.25            | 4.49            | 100.00           |
| $\emptyset$ | $0.99 \pm 0.01$ | $1.11 \pm 2.67$ | $0.01 \pm 0.04$ | $0.94 \pm 0.80$ | $97.51 \pm 3.09$ |

Tabelle 5.6.: Maximal erreichbare Segmentierungsgenauigkeit eines Radialstrahl-basierten Segmentierungsalgorithmus

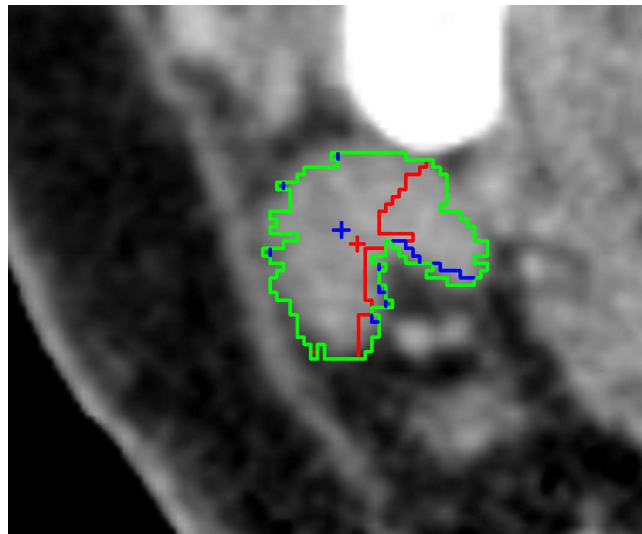


Abbildung 5.3.: Maximal erreichbare Segmentierungsgenauigkeit bei unterschiedlichen Saatpunkten im Vergleich zur Referenzsegmentierung (grün): Schwerpunkt (rot) und bestmöglicher Saatpunkt (blau)

ohne dabei die Oberfläche an anderer Stelle zu schneiden. Dies begrenzt die Menge der Formen, die segmentiert werden können, auf sternkonvexe Formen. Inwiefern diese Annahme auf Lymphknoten zutrifft, soll nun in diesem Abschnitt geklärt werden. Dazu wird für jedes zur Referenzsegmentierung dazugehörigen Voxel überprüft, ob dieses eine direkte Verbindung zum Schwerpunkt des Lymphknotens hat, ohne dabei jemals die Segmentierung zu verlassen. Ausgehend davon wird das bestmögliche Segmentierungsvolumen erstellt, indem diejenigen Voxel gesetzt werden, welche dem beschriebenen Kriterium genügen. Die in Abschnitt 5.1.1 beschriebenen Fehler-/Ähnlichkeitsmaße wurden dann zwischen diesen bestmöglichen Segmentierungen und den Referenzsegmentierungen berechnet. Statistiken über die Ergebnisse sind in Tabelle 5.6 zu finden.

Es zeigt sich, dass in den allermeisten Fällen die Form, bis auf einige wenige Voxel am Lymphknotenrand, komplett durch eine Radialstrahlenrepräsentation dargestellt werden kann und die Einschränkung auf Sternkonvexität ausgehend vom Schwerpunkt keinen Schwachpunkt des Segmentierungsverfahrens darstellt.

Der schlechteste Wert von  $\mathcal{G} = 79.79$  ist eine absolute Ausnahme und dennoch deutlich höher als die durchschnittliche durch manuelle Segmentierung erreichbare Segmentierungsgüte von  $\mathcal{G} = 75$ . Dies liegt an der speziellen Form des Lymphknotens. Der Schwerpunkt der Referenzsegmentierung, welcher als Saatpunkt verwendet wurde, ist in diesem Fall ungeeignet (siehe Abbildung 5.3). Wird der Saatpunkt allerdings verschoben, kann im besten Fall eine Segmentierungsgüte von  $\mathcal{G} = 94.26$ , also sehr nahe perfekter Überlappung auch für diesen Lymphknoten erreicht werden. Wird bei der manuellen Platzierung des Saatpunkts durch einen Anwender auf diese Limitierung des Verfahrens geachtet, kann also die Menge der unterstützten Formen im Vergleich zur Verwendung des Objektschwerpunktes als Saatpunkt weiter erhöht werden, sodass auch sehr ungewöhnliche Lymphknotenformen unterstützt werden.

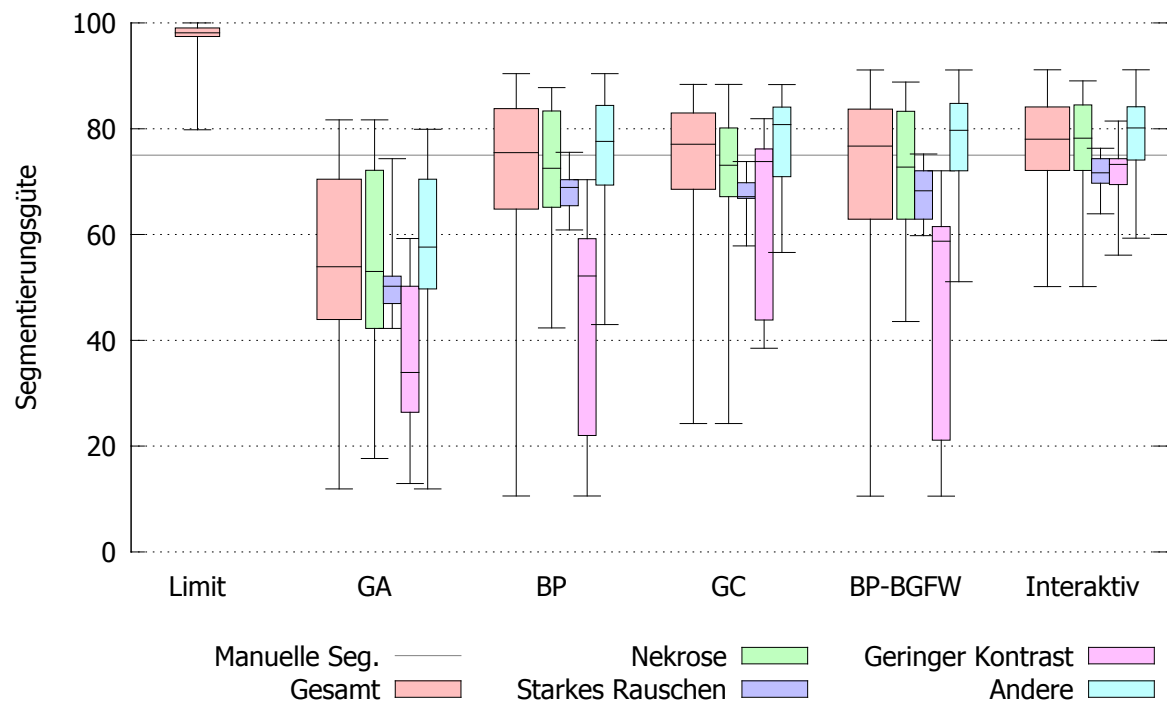


Abbildung 5.4.: Minimum, 0.25-Quantil, Median, 0.75-Quantil und Maximum der Segmentierungsgüte im Vergleich für unterschiedliche Verfahren und unterschiedliche Typen von Lymphknoten

**Segmentierungsgenauigkeit:** Da im vorangegangenen Abschnitt die theoretische Anwendbarkeit von Radialstrahl-basierten Segmentierungsmethoden auf Lymphknoten gezeigt wurde, widmet sich dieser Abschnitt der wohl wichtigsten Eigenschaft: der Segmentierungsgüte. Gegenübergestellt werden dabei die in Abschnitt 4.1.4 beschriebenen Optimierungsverfahren Gradientenabstieg (GA), Belief-Propagation (BP) und Graph-Cut (GC) mit der jeweils am besten geeigneten Kombination aus bildgradienten- und intensitätsbasierten Kostenfunktionen sowie den dazugehörigen optimalen Parametern, auf deren Bestimmung anhand der Trainingsmengen in Abschnitt 4.1.7 ausführlich eingegangen wurde.

In einem ersten Experiment wurden die Verfahren ohne die in Abschnitt 4.1.8 beschriebenen Erweiterungen auf die Lymphknoten der Testmengen angewendet. Um unabhängige und damit aussagekräftige Ergebnisse zu erhalten, wurde darauf geachtet, dass der zu segmentierende Lymphknoten sich nicht in derjenigen Trainingsmenge befand, aus der die Parameter des Segmentierungsverfahrens bestimmt wurden. Der Saatpunkt wurde in den Schwerpunkt der Referenzsegmentierung gesetzt. Anschließend wurden die in Abschnitt 5.1.1 beschriebenen Ähnlichkeits- bzw. Fehlermaße zwischen dem jeweiligen Segmentierungsergebnis und der Referenzsegmentierung berechnet. Statistiken über die Ergebnisse der verschiedenen Maße sind in Tabelle 5.7 eingetragen. Ausgewählte Segmentierungsergebnisse sind in Abbildung 5.5 dargestellt und ein graphischer Vergleich der Segmentierungsgüte  $\mathcal{G}$  ist in Abbildung 5.6 für verschiedene Klassen von Lymphknoten zu finden.

Es zeigt sich, dass globale Optimierungsverfahren deutlich bessere Ergebnisse als das lokale Optimie-

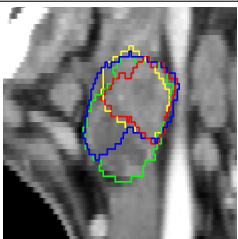
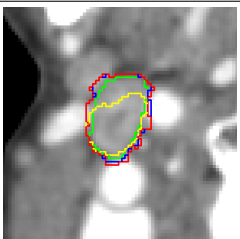
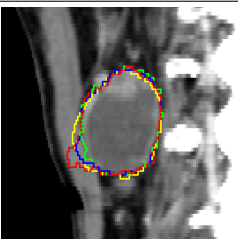
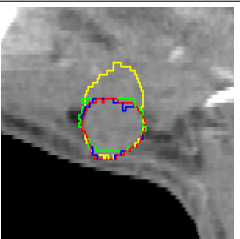
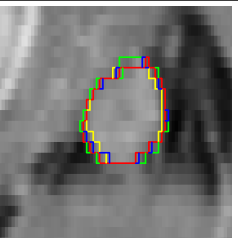
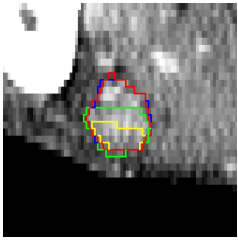
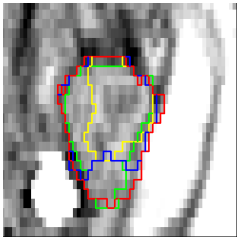
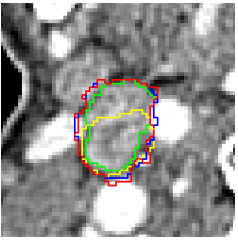
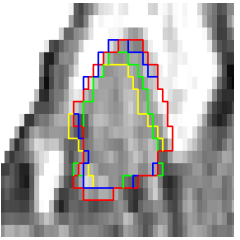
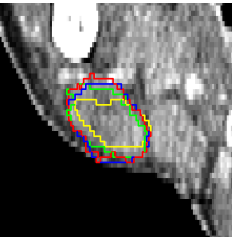
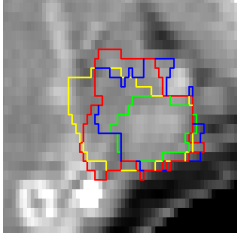
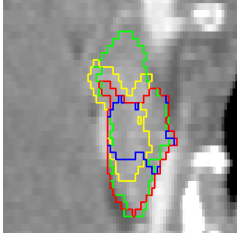
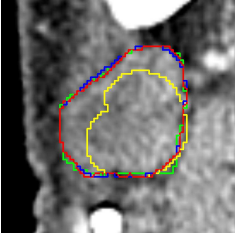
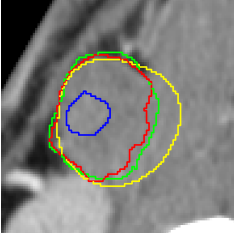
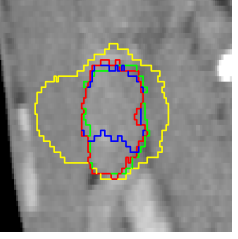
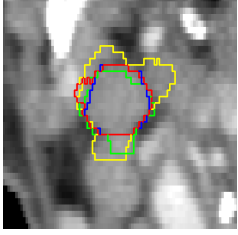
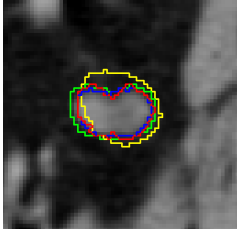
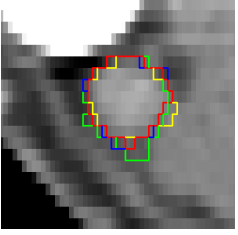
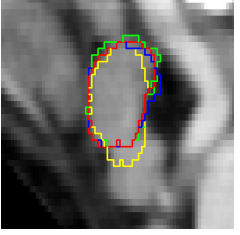
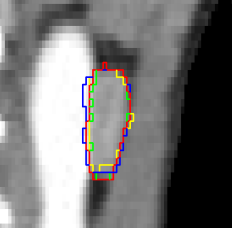
|               | Minimum  | 0.25-Quant.   | Median   | 0.75-Quant   | Maximum  |
|---------------|--|---|--|--|--|
| Nekrose       |   |   |   |   |   |
| Stark Rausch. |   |   |   |   |   |
| Gering Kontr. |   |   |   |   |   |
| Andere        |  |  |  |  |  |

Abbildung 5.5.: Beispielergebnisse des auf GA (gelb), BP (blau) und GC (rot) basierenden Segmentierungsverfahrens im Vergleich zur Referenzsegmentierung (grün) für unterschiedliche Lymphknotentypen. Die Beispiele wurden anhand der erreichten Segmentierungsgüte des GC Verfahrens ausgewählt.

|             | DSC  |      |      | VOE (%) |       |       | ASD (mm) |      |      | MSD (mm) |       |       | $\mathcal{G}$ |       |       |
|-------------|------|------|------|---------|-------|-------|----------|------|------|----------|-------|-------|---------------|-------|-------|
|             | GA   | BP   | GC   | GA      | BP    | GC    | GA       | BP   | GC   | GA       | BP    | GC    | GA            | BP    | GC    |
| Minimum     | 0.35 | 0.22 | 0.55 | 19.48   | 11.18 | 13.02 | 0.26     | 0.11 | 0.14 | 1.44     | 1.35  | 1.33  | 11.29         | 10.58 | 24.27 |
| 0.25-Quant. | 0.58 | 0.77 | 0.76 | 35.09   | 19.41 | 20.36 | 0.57     | 0.26 | 0.26 | 3.12     | 2.08  | 2.07  | 43.39         | 64.80 | 68.55 |
| Median      | 0.71 | 0.83 | 0.86 | 44.40   | 28.50 | 24.47 | 0.89     | 0.39 | 0.36 | 4.28     | 3.06  | 3.00  | 53.91         | 75.48 | 77.06 |
| 0.75-Quant. | 0.79 | 0.89 | 0.89 | 58.76   | 37.12 | 38.33 | 1.17     | 0.65 | 0.53 | 5.56     | 4.93  | 4.45  | 70.45         | 83.80 | 82.97 |
| Maximum     | 0.89 | 0.94 | 0.93 | 78.91   | 87.52 | 61.80 | 3.49     | 3.11 | 2.04 | 12.98    | 11.68 | 12.78 | 81.67         | 90.39 | 88.37 |
| Ø           | 0.68 | 0.81 | 0.83 | 47.07   | 29.94 | 28.61 | 0.99     | 0.55 | 0.46 | 5.02     | 3.99  | 3.65  | 54.15         | 70.95 | 73.24 |
| Std. Abw.   | 0.14 | 0.12 | 0.09 | 15.81   | 14.03 | 11.65 | 0.65     | 0.50 | 0.34 | 2.88     | 2.75  | 2.42  | 18.68         | 16.50 | 13.49 |

Tabelle 5.7.: Statistiken über verschiedene Fehler-/Ähnlichkeitsmäße von Segmentierungsergebnissen der in dieser Arbeit entwickelten Segmentierungsalgorithmen im Vergleich zur Referenzsegmentierung auf 49 Lymphknoten.

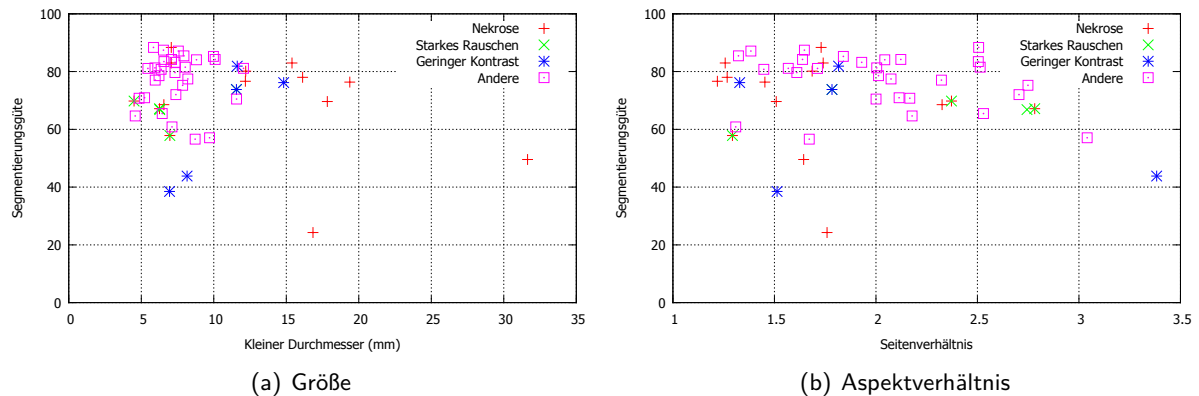


Abbildung 5.6.: Segmentierungsgüte des GC-Verfahrens in Abhängigkeit verschiedener Eigenschaften von Lymphknoten

runungsverfahren erzielen, wobei das auf Graph-Cut basierende Verfahren in der Regel leicht besser als das auf Belief-Propagation basierende Verfahren abschneidet. Im Allgemeinen werden bei Lymphknoten mit zentralem nekrotischen Gewebe ähnlich gute Ergebnisse wie bei anderen Lymphknoten erzielt. Starkes Rauschen führt zu einem geringen Verlust an Segmentierungsgenauigkeit. Der größte Abstrich an Genauigkeit kommt durch geringen Kontrast an Teilen der Lymphknotengrenze zustande. Dies ist zu erwarten, da in homogenen Regionen selbst ein erfahrener Spezialist die Lymphknotengrenze lediglich erahnen kann.

Zur Untersuchung des Einflusses unterschiedlicher Lymphknoteneigenschaften auf die Segmentierungsgüte wurden für das beste in dieser Arbeit vorgestellte Verfahren – das Graph-Cut basierte Verfahren – weitere Experimente durchgeführt. Ausgehend von den Referenzsegmentierungen wurden zunächst die Dimensionen der umrandenden Quader mit minimalem Volumen berechnet. Die kleinste Dimension (Durchmesser) beschreibt die Größe des Lymphknotens. Das Verhältnis zwischen größter und kleinster Dimension wird als Aspektverhältnis eines Lymphknotens bezeichnet. Es gibt an, wie länglich ein Lymphknoten ist und beschreibt dadurch unterschiedliche Formen. Die erzielte Segmentierungsgüte wurde nun in Abhängigkeit der Größe (Abbildung 5.6(a)) bzw. des Aspektverhältnisses (Abbildung 5.6(b)) aufgetragen. In beiden Abbildung ist keine klare Abhängigkeit zwischen den Eigenschaften und der erzielten Segmentierungsgüte zu erkennen. Daraus kann geschlossen werden, dass das Verfahren sowohl für alle auftretenden Größen als auch für unterschiedliche Formen gleichermaßen geeignet ist.

**Robustheit:** Neben einer Arbeitsreduktion für den Radiologen ist der Hauptvorteil von automatischen Segmentierungsverfahren, dass das Ergebnis bei gleichen Eingabedaten immer identisch ist. Im Gegensatz dazu stellt die überlicherweise sehr hohe Inter- aber auch Intra-Beobachter-Variabilität von manueller Segmentierung ein großes Problem dar. Insbesondere wenn Eigenschaften von Objekten innerhalb einer Zeitserie verglichen werden sollen, ist eine konsistente Segmentierung der Objekte von größter Bedeutung und zum Teil noch wichtiger als die erzielte Segmentierungsgenauigkeit.

Durch die manuelle Vorgabe des Saatpunkts handelt es sich bei dem in dieser Arbeit vorgestellten Segmentierungsverfahren um ein semi-automatisches Verfahren. Das Segmentierungsergebnis

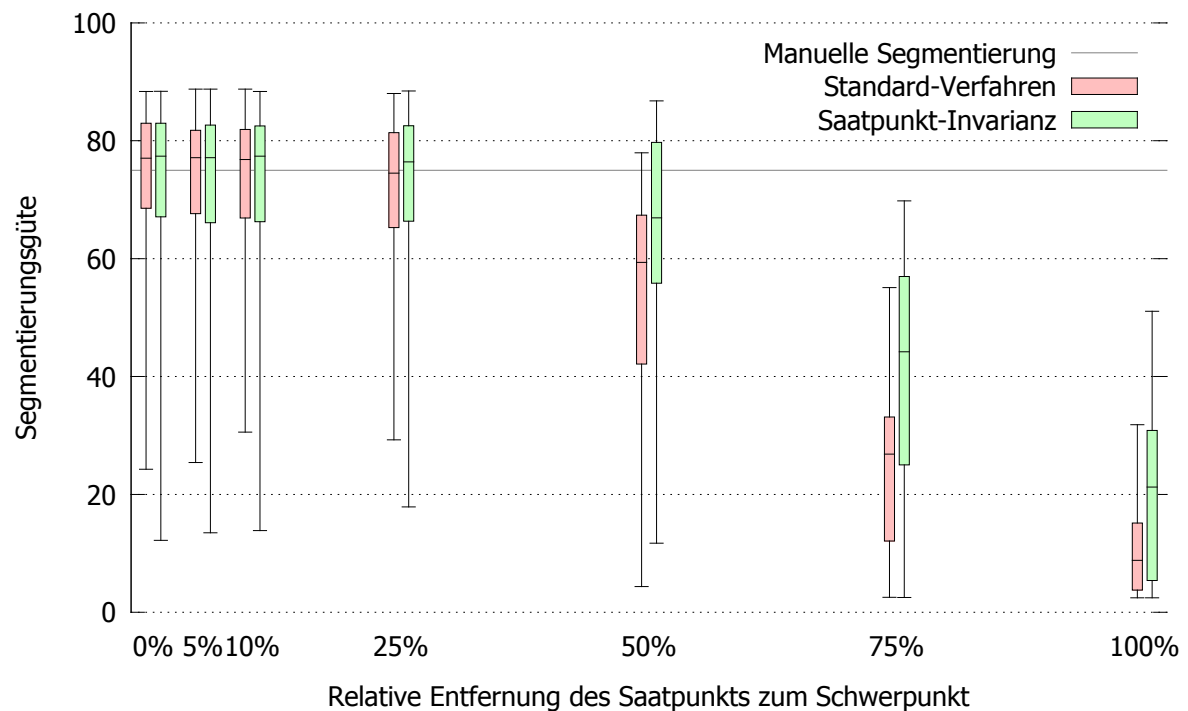


Abbildung 5.7.: Minimum, 0.25-Quantil, Median, 0.75-Quantil und Maximum der durchschnittlichen Segmentierungsgüte  $\bar{g}$  für verschiedene Entfernungen des Saatpunkts zum Lymphknotenschwerpunkt

ist prinzipbedingt von der Wahl des Saatpunkts abhängig. Diese Abhängigkeit soll nun untersucht werden.

Es ist zu erwarten, dass mit steigendem Abstand des Saatpunkts zum Lymphknotenschwerpunkt die Segmentierungsgenauigkeit tendenziell abnimmt. Wie stark dieser Effekt zum Tragen kommt, soll nun geklärt werden. Zusätzlich zu dem auf Graph-Cut basierenden Verfahren wird auch die in Abschnitt 4.1.8 erläuterte saatpunktinvariante Erweiterung untersucht. Für jeden der 49 Lymphknoten wurden zunächst jeweils 66 gleichverteilte Punkte mit konstantem relativen Abstand zum Lymphknotenschwerpunkt bestimmt. Der Abstand betrug 5, 10, 25, 50, 75 bzw. 100% der Größe des Lymphknotens, welche durch den kleinsten Durchmesser gegeben ist. Diese Punkte wurden dann als Saatpunkte für die beiden Segmentierungsverfahren verwendet. Für jeden Lymphknoten und jeden relativen Abstand wurden die 66 im Allgemeinen unterschiedlichen Segmentierungsgüten  $g$  im Vergleich zur Referenzsegmentierung gemessen sowie der Durchschnitt und die Standardabweichung bestimmt. Statistiken dieser Werte für alle Lymphknoten sind in Tabelle 5.7 bzw. Tabelle 5.8 zu sehen.

Bzgl. der durchschnittlichen Segmentierungsgüte ist ein starker Abfall für alle relativen Entfernungen zum Schwerpunkt größer als 25% zu erkennen. Dies trifft sowohl auf das Standardverfahren als auch auf das saatpunktinvariante Verfahren zu. Wohingegen die Werte für relative Entfernungen bis zu 25% bei beiden Verfahren vergleichbar sind, erzielt das saatpunktinvariante Verfahren für größere Entfernungen eine höhere durchschnittliche Segmentierungsgüte. Bei der Standardabweichung ist ein

## 5. Experimente & Ergebnisse

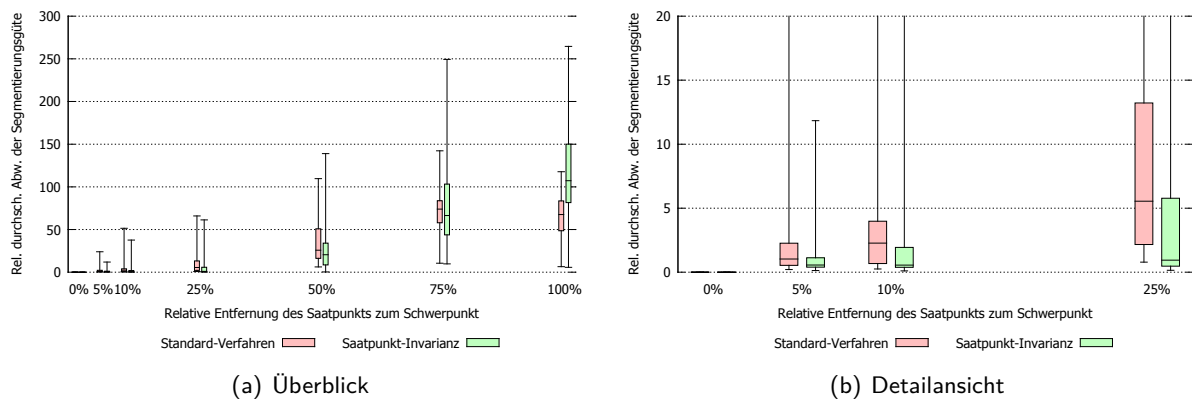


Abbildung 5.8.: Minimum, 0.25-Quantil, Median, 0.75-Quantil und Maximum der Standardabweichung der Segmentierungsgüte  $\bar{G}$  für verschiedene Entfernungen des Saatpunkts zum Lymphknotenschwerpunkt

Vorteil der Erweiterung schon bei geringen Entfernungen erkennbar. So ist der Median der Standardabweichung des saarpunktinvarianten Verfahrens bei 25% relativer Entfernung mit 0.56 nahezu identisch zu dem bei 5% (0.43). Das Standardverfahren hingegen weist bei der sehr geringen relativen Entfernung von 5% schon eine Median der Standardabweichung von 0.74 und bei 25% bereits einen Wert von 4.05 auf. Die gewünschte Erhöhung der Robustheit durch das saarpunktinvariante Verfahren konnte dadurch nachgewiesen werden.

Ausgehend davon kann nun eine der Kernfragen dieser Arbeit beantwortet werden: Kann das automatische Verfahren eine größere Segmentierungskonsistenz als manuelle Segmentierung erzielen, und somit insbesondere bei der Zeitserienanalyse quantitativer Merkmale von Lymphknoten einen Mehrwert liefern? Dazu muss zunächst bekannt sein, welche relativen Entfernungen zum Lymphknotenschwerpunkt von menschlichen Beobachtern bei der Saarpunktwahl erreicht werden. Um dies zu bestimmen, wurde ein Radiologe gebeten in den ungefähren Schwerpunkt der 49 Lymphknoten zu klicken. Von diesen Punkten ausgehend wurde dann die Entfernung zum Schwerpunkt relativ zur Lymphknotengröße gemessen. Diese betrug durchschnittlich 21.6% bei einem Median von 15.7% und ist somit deutlich innerhalb der 25%-Grenze.

Im Folgenden soll nun auf ein bestimmtes oft verwendetes Merkmal, das Volumen, eingegangen werden. Grundsätzlich sollte die Segmentierungsgenauigkeit so groß wie möglich und bei ein und demselben Lymphknoten die Standardabweichung des Merkmals (hier: Volumen) bei unterschiedlichen Eingabedaten (hier: Saarpunkt) so gering wie möglich sein. Bei dem weiter oben beschriebenen Experiment wurden neben dem Durchschnitt und Standardabweichung der Segmentierungsgüte für verschiedene relative Entfernungen des Saatpunkts zum Lymphknotenschwerpunkt auch der Durchschnitt und die Standardabweichung des erzielten Lymphknotenvolumens bestimmt. Der Durchschnitt und die Standardabweichung der durchschnittlichen Segmentierungsgüte und Standardabweichung des erzielten Volumens sind für beide Verfahren in Abbildung 5.9 eingetragen. Zusätzlich wurden diese Werte auch für die manuelle Segmentierung aus den 5 vorliegenden Expertensegmentierungen bestimmt und ebenfalls dargestellt.

Es zeigt sich, dass bis zu einer relativen Entfernung zum Schwerpunkt von 25% der Lymphknoten-



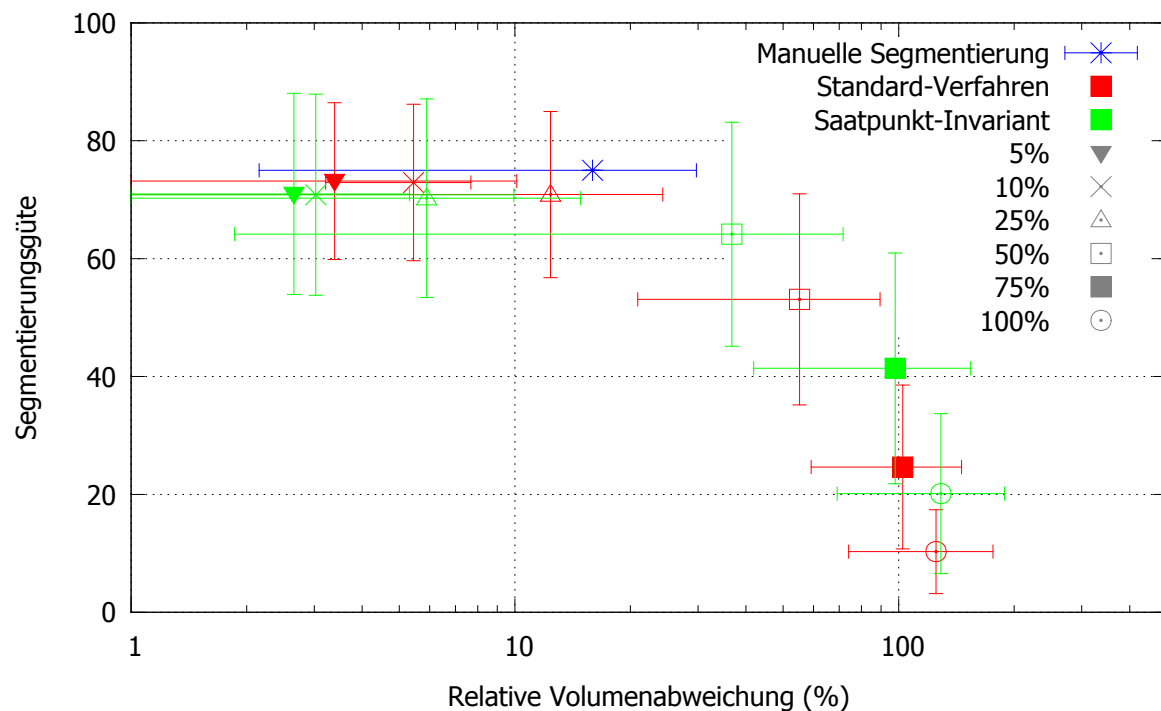


Abbildung 5.9.: Durchschnitt und Standardabweichung der durchschnittlichen Segmentierungsgüte sowie der relativen Volumenabweichung für unterschiedliche Entfernungen des Saatpunkts zum Lymphknotenschwerpunkt. Bis zu einer relativen Entfernungen von 25% ist das saatpunktinvariante Verfahren robuster als manuelle Konturierung.

größe die durchschnittliche Segmentierungsgenauigkeit etwas geringer als die der manuellen Segmentierung ist. Die Standardabweichung des erzielten Volumens ist allerdings deutlich geringer. Das saatpunktinvariante Verfahren schneidet deutlich besser als das Standard-Verfahren ab. Bei 25% relativer Entfernung, also etwas höher als in der Praxis zu erwarten ist, beträgt die durchschnittliche Volumenstandardabweichung 5.9%. Bei der manuellen Segmentierung ist diese mit 15.9% fast drei mal so hoch. Der Vergleich von Volumen mit Hilfe des in dieser Arbeit entwickelten Lymphknotensegmentierungsverfahrens ist also präziser als eine manuelle Beurteilung durch Experten in der Praxis.

**Laufzeit:** Bei vollautomatischen Segmentierungsverfahren spielt die Laufzeit eine eher untergeordnete Rolle, weil die Segmentierung im Hintergrund automatisiert gleich nach der Bildakquise ablaufen kann und somit die Zeit, die bis zur Begutachtung durch einen Experten verstreicht, genutzt werden kann. Bei semi-automatischen Verfahren, wie die in dieser Arbeit zur Lymphknotensegmentierung vorgestellten Verfahren, ist dies allerdings nicht möglich, weil eine Anwendereingabe – das manuelle Setzen des Saatpunkts – notwendig ist. Daraufhin erwartet der Anwender das Ergebnis innerhalb weniger Sekunden, um aus dem Segmentierungsergebnis und evtl. extrahierten Merkmalen, wie z.B. dem Volumen, Schlüsse ziehen zu können.

Prinzipbedingt ist das lokale Gradientenabstiegsverfahren deutlich weniger komplex als die beiden

## 5. Experimente & Ergebnisse

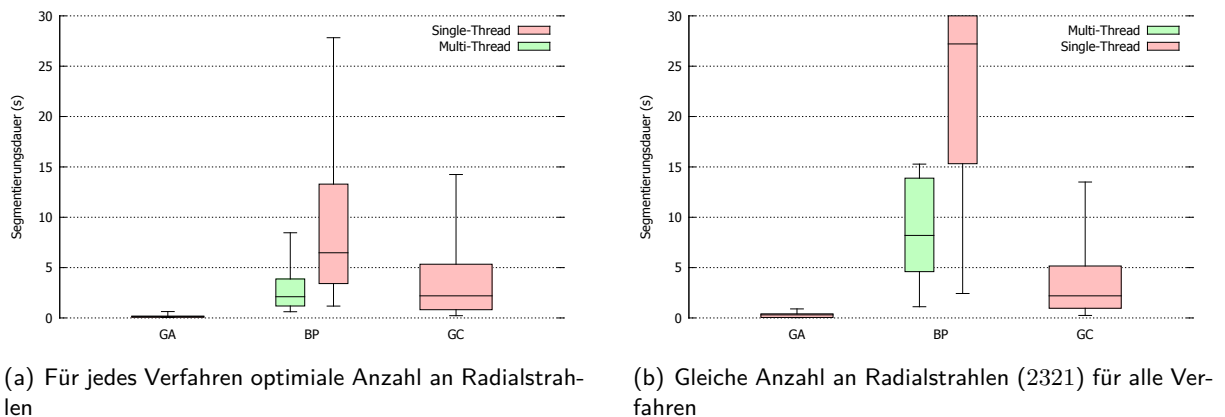


Abbildung 5.10.: Vergleich der Laufzeit der unterschiedlichen Optimierungsverfahren

| Laufzeit (s) | GA   |       | BP (single-thread) |       | BP (multi-thrad) |       | GC    |       |
|--------------|------|-------|--------------------|-------|------------------|-------|-------|-------|
|              | opt. | norm. | opt.               | norm. | opt.             | norm. | opt.  | norm. |
| Minimum      | 0.05 | 0.03  | 1.18               | 2.43  | 0.61             | 1.11  | 0.23  | 0.25  |
| 0.25-Quant   | 0.07 | 0.05  | 3.41               | 15.32 | 1.19             | 4.60  | 0.82  | 0.96  |
| Median       | 0.11 | 0.31  | 6.47               | 27.21 | 2.11             | 8.20  | 2.20  | 2.19  |
| 0.75-Quant   | 0.18 | 0.41  | 13.29              | 50.26 | 3.87             | 13.88 | 5.33  | 5.16  |
| Maximum      | 0.63 | 0.90  | 27.83              | 53.72 | 8.46             | 15.27 | 14.24 | 13.46 |
| Ø            | 0.14 | 0.28  | 8.50               | 29.65 | 2.68             | 8.69  | 3.67  | 3.74  |
| Std. Abw.    | 0.10 | 0.19  | 6.27               | 16.75 | 1.82             | 4.60  | 3.62  | 3.64  |

Tabelle 5.8.: Vergleich der Laufzeit der unterschiedlichen Optimierungsverfahren mit optimierten bzw. normalisierten Parametern. Alle Werte sind in Sekunden angegeben.

globalen Verfahren, jedoch kann zumindest bei dem auf Belief-Propagation basierenden Verfahren durch Parallelisierung eine deutliche Reduktion der Laufzeit erreicht werden. Im Folgenden soll die Laufzeit von praktischen Implementierungen der einzelnen Verfahren quantitativ untersucht werden. Durchgeführt wurden die Experimente auf einem Intel Core i7-2600K Prozessor im 3.40 GHz und 8 Prozessorkernen. Da das Gradientenabstiegsverfahren keine große Komplexität besitzt, wurde bei der Implementierung auf die Verwendung von Parallelisierung verzichtet. Auch bei Graph-Cut wird lediglich ein Prozessorkern verwendet. Es kam die Implementierung von Boykov und Kolmogorov [141] zum Einsatz. Bei der eigenen Implementierung von Belief-Propagation wurde hingegen optional Parallelisierung innerhalb einer Iteration bei der Berechnung der Nachrichten verwendet.

Mit allen drei Implementierungen wurden alle 49 Lymphknoten mit den optimalen Parametern segmentiert. Dabei wurde die Laufzeit gemessen. Da die Laufzeit allerdings sehr stark von der Anzahl der Strahlen abhängt, haben die Ergebnisse in der Praxis zwar eine große Aussagekraft, ein eher theoretischer Vergleich ist allerdings nicht möglich, weil die Anzahl der Strahlen für alle Verfahren unterschiedlich ist. Daher wurden in einem weiteren Experiment alle Verfahren mit derselben normalisierten Strahlenanzahl von  $n = 2321$  auf alle 49 Lymphknoten angewendet und die Laufzeit gemessen.

Statistiken über die Ergebnisse sind in Tabelle 5.8 eingetragen und in Abbildung 5.10 dargestellt.

Das lokale Gradientenabstiegsverfahren braucht für die Segmentierung eines Lymphknotens deutlich unter einer Sekunde und ist damit problemlos im klinischen Alltag einsetzbar. Die globalen Verfahren brauchen mit durchschnittlich 2.68 Sekunden für Belief-Propagation und 3.67 Sekunden für Graph-Cut deutlich länger. Diese Werte befinden sich aber dennoch in einem für die klinische Einsetzbarkeit akzeptablen Bereich. In einigen Fällen erhöht sich die Segmentierungsdauer von Graph-Cut auf bis zu 14.24 Sekunden. Bei Belief-Propagation ist die Verwendung von Parallelisierung dringend notwendig. Wird darauf verzichtet, steigt die Segmentierungsdauer auf durchschnittlich 8.50 Sekunden, im schlimmsten Fall sogar auf 27.83 Sekunden.

Werden die Segmentierungsverfahren mit gleicher Strahlenanzahl angewendet, ergibt sich ein Vorteil (durchschnittlich 3.74 Sekunden) von Graph-Cut gegenüber Belief-Propagation (8.50 Sekunden). Weitere Parallelisierung, bspw. durch eine GPU-Implementierung, kann diesen Unterschied jedoch vermutlich aufheben. Das lokale Gradientenabstiegsverfahren ist mit durchschnittlich 0.28 Sekunden für die Segmentierung eines Lymphknotens deutlich am schnellsten.

**Bildgradientenbasiertes Formwissen:** Das Ziel der Entwicklung des bildgradientenbasierten Formwissens (BGFW) war Lymphknoten, welche nicht sonderlich kugelförmig sind, besser zu unterstützen. Das normalerweise verwendete Formwissen verlangt hingegen, dass benachbarte Strahlen einen möglichst ähnlichen Radius haben und präferiert somit Kugeln. Das auf Belief-Propagation basierende Verfahren wurde dazu unter Verwendung des bildgradientenbasierten Formwissens auf die Lymphknoten der Testmengen angewendet. Die erhaltenen Segmentierungen wurden anschließend mit den Referenzsegmentierungen verglichen. Die Ergebnisse sind in Tabelle 5.9 zu sehen und in Abbildung 5.4 eingetragen. Einige Beispielsegmentierungsergebnisse sind in Abbildung 5.11 zu sehen.

Durchschnittlich verbessert sich die Segmentierungsgüte gegenüber dem auf Belief-Propagation basierenden Verfahren mit dem traditionell verwendeten Formwissen um 1%. Das Verfahren kann aber dennoch nicht mit dem Graph-Cut basierenden Verfahren mithalten.

Bei diesem geringen Gewinn an Segmentierungsgüte stellt sich nun die Frage, ob bei weniger runden Lymphknoten tatsächlich ein größerer Gewinn an Segmentierungsgüte erreicht wird. Bei dem Beispiel in Abbildung 5.11(e) ist dies tatsächlich der Fall. Auf der rechten Seite kann sich das Verfahren mit BGFW besser an die konkave Form anpassen. Doch wie ist das im Allgemeinen? In Abbildung 5.12 wurde dazu der Gewinn an Segmentierungsgüte gegenüber dem Aspektverhältnis aus größtem und kleinstem Durchmesser, einem Maß für die Abweichung von runden Formen, aufgetragen. Sowohl Verluste als auch Gewinne sind zu erkennen. Die in dieser Abbildung dargestellte Trendlinie bestätigt durch den Anstieg hin zu größeren Aspektverhältnissen die Entwicklungsabsicht, weniger runde Formen besser zu unterstützen, wobei der Effekt auch nur gering ausfällt.

**Interaktive Segmentierung:** Da bei einigen Lymphknoten, insbesondere mit geringem Bildkontrast an der Lymphknotengrenze, mit den semi-automatischen Verfahren keine zufriedenstellenden Ergebnisse erzielt werden konnten, wurde in dieser Arbeit ein interaktives Verfahren zum manuellen Eingriff während des Segmentierungsprozesses vorgestellt. Ein Anwender segmentierte damit alle 49 Lymphknoten. Dabei wurde neben der Segmentierungsdauer auch gemessen, wie oft der Anwender in den Segmentierungsprozess eingegriffen hat. Anschließend wurde die Segmentierungsgenauigkeit wie üblich im Vergleich zur Referenzsegmentierung gemessen. Ergebnisse über die erreichte Segmen-

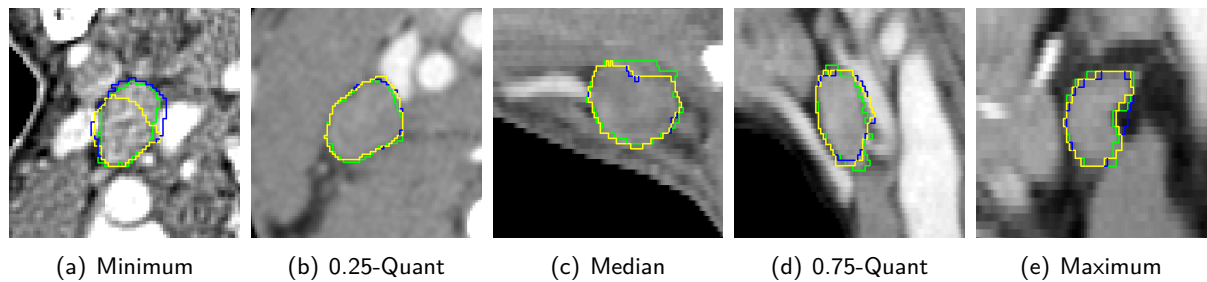


Abbildung 5.11.: Beispielergebnisse des bildgradientenbasierten Formwissens (gelb) in dem auf Belief-Propagation basierenden Segmentierungsverfahren im Vergleich zum Standardformwissen (blau) und der Referenzsegmentierung (grün). Die Auswahl der Beispiele erfolgte anhand des Gewinns an Segmentierungsgüte des BGFWs.

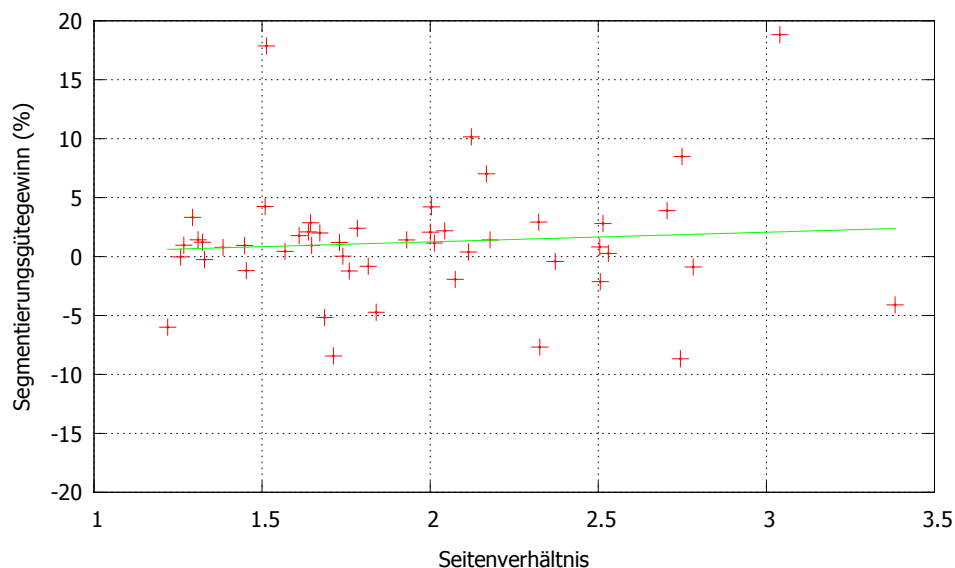


Abbildung 5.12.: Das bildgradientenbasierte Formwissen erzielt bei Lymphknoten mit großem Aspektverhältnis tendenziell einen größeren Gewinn an Segmentierungsgenauigkeit.

|             | DSC  | VOE (%) | ASD (mm) | MSD (mm) | $\mathcal{G}$ |
|-------------|------|---------|----------|----------|---------------|
| Minimum     | 0.22 | 11.30   | 0.10     | 1.20     | 10.55 −0.3%   |
| 0.25-Quant. | 0.77 | 18.69   | 0.26     | 2.01     | 62.89 −2.9%   |
| Median      | 0.85 | 26.47   | 0.35     | 3.13     | 76.73 +1.7%   |
| 0.75-Quant. | 0.90 | 36.80   | 0.68     | 5.00     | 83.70 −0.1%   |
| Maximum     | 0.94 | 87.64   | 3.12     | 11.68    | 91.10 +0.8%   |
| Ø           | 0.82 | 29.21   | 0.53     | 3.96     | 71.67 +1.0%   |
| Std. Abw.   | 0.12 | 14.02   | 0.50     | 2.70     | 16.40 −0.6%   |

Tabelle 5.9.: Segmentierungsgenauigkeit des bildgradientenbasierten Formwissens in dem auf Belief-Propagation basierenden Segmentierungsverfahren und der Änderung gegenüber dem Standardformwissen

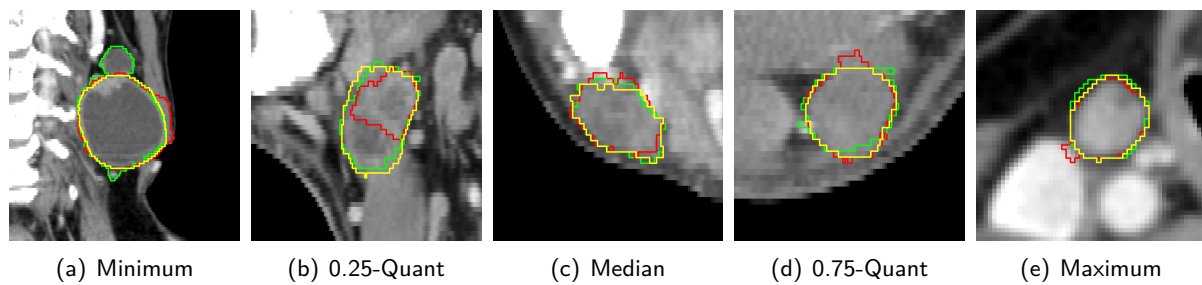


Abbildung 5.13.: Beispielergebnisse des interaktiven Segmentierungsverfahrens (gelb) im Vergleich zum semi-automatischen auf GC basierenden Verfahren (rot) und der Referenzsegmentierung (grün). Die Auswahl der Beispiele erfolgte anhand der erzielten Segmentierungsgüte des interaktiven Verfahrens.

|             | DSC  | VOE   | ASD  | MSD   | $\mathcal{G}$ | $\mathcal{G}$ ohne Interaktion | Dauer  | Eingriffe |
|-------------|------|-------|------|-------|---------------|--------------------------------|--------|-----------|
| Minimum     | 0.70 | 8.95  | 0.10 | 1.20  | 50.15         | 10.58                          | 1.9 s  | 0         |
| 0.25-Quant. | 0.82 | 19.16 | 0.25 | 2.00  | 72.11         | 64.80                          | 4.5 s  | 0         |
| Median      | 0.86 | 24.13 | 0.33 | 2.55  | 78.02         | 75.48                          | 11.6 s | 1         |
| 0.75-Quant. | 0.89 | 30.00 | 0.50 | 3.55  | 84.11         | 83.80                          | 17.1 s | 2         |
| Maximum     | 0.95 | 45.95 | 1.17 | 12.04 | 91.12         | 90.39                          | 35.5 s | 5         |
| $\emptyset$ | 0.85 | 25.23 | 0.38 | 3.09  | 77.13         | 70.95                          | 12.1 s | 1.31      |
| Std. Abw.   | 0.06 | 8.98  | 0.20 | 1.78  | 8.96          | 16.50                          | 7.8 s  | 1.37      |

Tabelle 5.10.: Segmentierungsgenauigkeit (auch im Vergleich zum automatischen BP-basierten Verfahren), Dauer und Anwenderinteraktionsmenge des interaktiven Segmentierungsverfahrens

tierungsgenauigkeit, Dauer und die Menge der benötigten Anwenderinteraktion ist in Tabelle 5.10 eingetragen und einige Beispielsegmentierungsergebnisse sind in Abbildung 5.13 zu sehen.

In 39% der Fälle war der Anwender schon mit der vorgeschlagenen Segmentierung zufrieden, und es war überhaupt keine Interaktion nötig. In den anderen Fällen korrigierte der Anwender die Kontur an durchschnittlich 1.85 Stellen. Die Segmentierungsgenauigkeit hat sich dadurch erheblich verbessert, insbesondere bei denjenigen Lymphknoten, die mit dem semi-automatischen Verfahren schlecht segmentiert wurden. So verbessert sich die minimale Segmentierungsgüte um 107% von 24.27 auf 50.15. Auch das zuvor sehr schlechte Abschneiden von Lymphknoten mit geringem Bildkontrast konnte dadurch weitestgehend in der Griff bekommen werden. Insgesamt ist die erzielte durchschnittliche Segmentierungsgüte mit 77.13 sogar um ca. 3% höher als die der manuellen Segmentierung. Ein Grund dafür könnte sein, dass die Segmentierungsdauer deutlich geringer als bei der schichtweisen manuellen Segmentierung ist und somit die Ermüdung des Anwenders geringer sein könnte. Des weiteren handelt es sich beim interaktiven Verfahren um ein echtes 3D-Verfahren. Damit ist die Segmentierung über die Schichten hinweg konsistent und auch in Schichten mit geringem Bildkontrast können durch den Einfluss der benachbarten Schichten sinnvolle Ergebnisse erzielt werden.

Die durchschnittliche Segmentierungsdauer von 12.1 Sekunden für jeden Lymphknoten sollte kein Problem in der klinischen Praxis sein, zumindest solange nur einige wenige Lymphknoten unter-

|             | Diese Arbeit | Region-Growing | EMSM  | Wang  | Barbu |
|-------------|--------------|----------------|-------|-------|-------|
| Minimum     | 24.27        | -              | 3.177 | 2.54  | 3.57  |
| 0.25-Quant. | 68.55        | -              | 41.27 | 24.95 | 34.54 |
| Median      | 77.06        | -              | 57.19 | 67.75 | 48.53 |
| 0.75-Quant. | 82.97        | -              | 68.52 | 78.71 | 53.48 |
| Maximum     | 88.37        | -              | 79.68 | 84.64 | 78.26 |
| Ø           | 73.24        | 5.09           | 52.09 | 54.74 | 43.52 |
| Std. Abw.   | 13.49        | -              | 21.12 | 29.25 | 18.24 |

Tabelle 5.11.: Vergleich der erzielten Segmentierungsgüte des in dieser Arbeit entwickelten Verfahrens (GC) mit der Literatur

sucht werden, so wie dies im Anwendungsszenario dieser Arbeit – der Untersuchung von Kopf-Hals-Karzinomen – der Fall ist.

Ein Nachteil der interaktiven Erweiterung ist ganz klar, dass sich die Abhängigkeit vom Anwender vergrößert. Konnte beim semi-automatischen Verfahren, insbesondere durch die saatkpunktinvariante Erweiterung, diese Abhängigkeit gering gehalten werden, so ist diese beim interaktiven Verfahren deutlich näher an manueller Segmentierung, weil der Anwender die Kontur in sehr großem Maße beeinflussen kann.

**Gegenüberstellung von Lymphknotensegmentierungsverfahren:** Die erzielte Segmentierungsgenauigkeit der in dieser Arbeit vorgestellten Verfahren soll nun mit anderen Verfahren aus der Literatur verglichen werden. Neben dem im Rahmen dieser Arbeit entwickelten Radialstrahl-basierten Verfahren mit Graph-Cut Optimierung werden folgende Verfahren betrachtet:

- *Region-Growing (RG)*: Ausgehend von einem Saatpunkt werden alle diejenigen benachbarten Voxel zur Segmentierung gezählt, welche einem Ähnlichkeitskriterium genügen. Die Bildintensität wird als Gaußverteilung modelliert. Alle benachbarten Voxel, deren Intensitäten sich innerhalb eines bestimmten Konfidenzintervalls befinden, werden zu der Segmentierung gezählt. Die Parameter der Gaußverteilung werden in jeder Iteration zunächst aus einer Umgebung um den Saatpunkt und später aus allen zur Segmentierung dazugehörigen Voxel berechnet. Dieses Verfahren ist nicht lymphknotenspezifisch und soll dennoch in den Vergleich einbezogen werden, weil es sich um ein häufig verwendetes Standardverfahren der medizinischen Bildverarbeitung handelt.
- Größeninvariantes *Masse-Feder-Modell (EMSM)*: Das in Abschnitt 4.1.1 vorgestellte Lymphknotensegmentierungsverfahren, welches auf den Arbeiten von Dornheim et al. [80][83] aufbaut, soll als Vertreter der modellbasierten Verfahren mit globalem Formwissen in den Vergleich einbezogen werden.
- Das Radialstrahl-basierte Verfahren von Wang et al. [93] besitzt dieselbe zugrundeliegende Datenstruktur wie das in dieser Arbeit verwendete Verfahren und auch das Finden der Radien für die einzelnen Strahlen geschieht durch Graph-Cut. Der Unterschied besteht in der Kostenfunktion. Wang et al. setzen dabei auf eine Kombination aus einem Homogenitäts- und Kantenterm. Ein zusätzlicher Konsistenzterm sorgt dafür, dass Radien, die über die Strahlen hinweg kon-

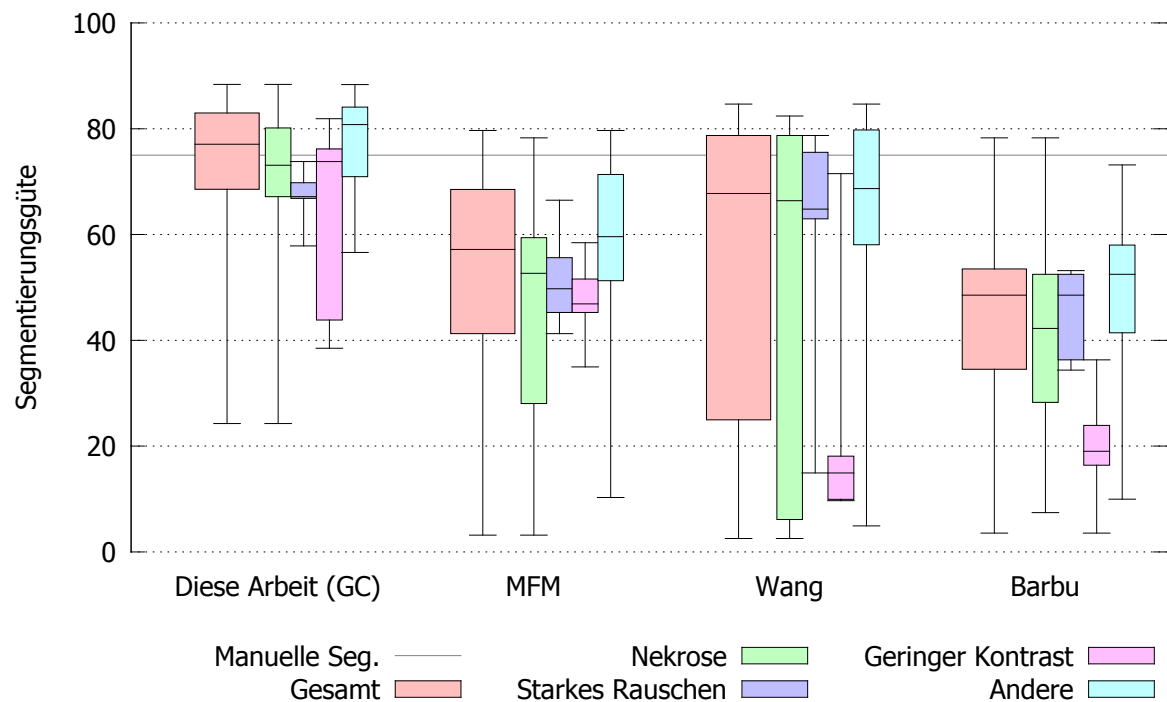


Abbildung 5.14.: Vergleich der Segmentierungsgüte des in dieser Arbeit vorgestellten Verfahrens mit der Literatur für unterschiedliche Typen von Lymphknoten. Jeder Balken stellt Minimum, 0.25-Quantil, Median, 0.75-Quantil und Maximum der Segmentierungsgüte dar.

sistent geringe Kantenkosten aufweisen, bevorzugt werden. Eine genauere Beschreibung ist in Abschnitt 3.1.5 zu finden.

- Auch das Verfahren von *Barbu* et al. [67] ist ein Radialstrahl-basiertes Verfahren, bei welchem ein Gradientenabstiegsverfahren zur Radienfindung eingesetzt wird. Die Kostenfunktion ist recht einfach. Entlang des Strahls wird ausgehend vom Saatpunkt, derjenige nächste Punkt gefunden, welcher um mindestens 50 HU von der Intensität des Saatpunkts abweicht. Die Kosten für einen beliebigen Punkt auf dem Strahl hängen dann von der Distanz zu diesem Punkt ab. Auch dieses Verfahren ist näher in Abschnitt 3.1.5 beschrieben.

Bei dem Region-Growing wurde eine Implementierung aus dem *Insight Toolkit* (ITK) verwendet und die Evaluation fand im Rahmen einer Masterarbeit [17] statt. Bei den anderen Verfahren handelt es sich um Eingenisimplementierungen. Die Parameter wurden aus den entsprechenden Arbeiten übernommen. Alle Verfahren wurden auf den Evaluationsdatensatz angewendet und mit den Referenzsegmentierungen verglichen. Die Ergebnisse sind in Tabelle 5.11 eingetragen und in Abbildung 5.14 dargestellt.

Es zeigt sich, dass die in dieser Arbeit vorgestellten Verfahren deutlich bessere Segmentierungsergebnisse als die in der Literatur beschriebenen Verfahren liefern. Die katastrophalen Ergebnisse des Region-Growings zeigen, dass solch einfache Standardverfahren nicht zur Segmentierung von Lymphknoten geeignet sind. Der Ansatz von Wang et al. liefert in vielen Fällen gute Ergebnisse, besonders

| Verfahren                      | #   | DSC  | VOE (%)   | ASD (mm)                                       | MSD (mm)  | Kommentar                  |
|--------------------------------|-----|--|---|--|---|----------------------------|
| Diese Arbeit<br>(Graph-Cut)    | 49  | 0.83 ± 0.09<br>0.76, 0.86, 0.89<br>0.55 – 0.93 | 28.61 ± 11.65<br>20.36, 24.47, 38.33<br>13.02 – 61.80 | 0.46 ± 0.34<br>0.26, 0.36, 0.53<br>0.14 – 2.04 | 3.65 ± 2.42<br>2.07, 3.00, 4.45<br>1.33 – 12.78 | 1-Punkt Initialisierung    |
| Diese Arbeit<br>(Belief-Prop.) | 49  | 0.85 ± 0.06<br>0.82, 0.86, 0.95<br>0.70 – 0.95 | 25.23 ± 8.89<br>19.16, 24.13, 30.00<br>8.95 – 45.95   | 0.38 ± 0.20<br>0.10, 0.25, 0.33<br>0.10 – 1.17 | 3.09 ± 1.78<br>2.00, 2.55, 3.55<br>1.20 – 12.04 | Interaktiv                 |
| Dornheim[81]                   | 40  | -  | 44.00 ± –   | 0.47 ± –                                       | 2.74 ± –  | 2-Punkt Initialisierung    |
|                                |     | -  | -   | 0.3 – 0.6                                      | 1.7 – 3.9                                       |                            |
| Maleike[79]                    | 29  | -  | -   | 0.4, 0.5, 0.6                                  | 0.9, 1.2, 3.8                                   | 1-Strich Initialisierung   |
| Maleike[79]                    | 29  | -  | -   | 0.4, 0.5, 0.6                                  | 1.0, 1.3, 2.3                                   | Interaktiv                 |
| Dornheim[82]                   | 40  | -  | 27.9 ± –  | 0.28 ± –                                       | 2.44 ± –  | 1-Punkt Initialisierung    |
| Wang[93]                       | 22  | 0.84 ± –                                       | -   | 0.47 ± 0.08                                    | 2.12 ± –  | 1-Punkt Initialisierung    |
| Barbu[67]                      | 308 | 0.80 ± 0.13                                    | -   | 0.83 ± 0.66                                    | –   | 1-Punkt Init., Achselhöhle |
| Barbu[67]                      | 455 | 0.76 ± 0.13                                    | -   | 1.0 ± 0.80                                     | –   | 1-Punkt, Becken&Abdomen    |

Tabelle 5.12.: Vergleich der in der Literatur angegebenen Ergebnisse (Durchschnitt ± Standardabweichung; 0.25-Quant, Median, 0.75-Quant; Minimum - Maximum)



für kleine, rundliche Lymphknoten. Allerdings zeigt der Ansatz deutliche Schwächen bei Lymphknoten mit geringem Bildkontrast und bei vergrößerten Lymphknoten mit nekrotischem Gewebe. Der Vorteil des Ansatzes von Barbu et al. besteht in der Einfachheit und der daraus resultierenden kurzen Laufzeit. Die Segmentierungsqualität ist aber im Allgemeinen nicht sonderlich hoch. Durch den Vergleich mit dem Masse-Feder-Modell kann geschlossen werden, dass auf Radialstrahlen basierende Verfahren mit globalem Optimierungsverfahren geeigneter als modellbasierte Verfahren zur Segmentierung von Lymphknoten sind. Der Grund besteht vermutlich darin, dass modellbasierte Verfahren üblicherweise eine lokale Anpassungsstrategie haben und das Segmentierungsergebnis sehr stark von der initialen Modellplatzierung abhängt. Durch die hohe Form- und Größenvarianz von Lymphknoten ist eine sinnvolle Initialisierung allerdings völlig unklar. Darüber hinaus scheint die Verwendung von globalem Formwissen, wie dies bei modellbasierten Verfahren der Fall ist, nicht geeignet zu sein, weil Lymphknoten keine spezielle globale Form bevorzugen.

Die Ergebnisse der selbst durchgeführten Experimente sollen nun der Vollständigkeit halber den in der Literatur erzielten Ergebnissen gegenübergestellt werden. Dazu wurden alle in der Literatur berichteten Ergebnisse in Tabelle 5.12 eingetragen. Die angegebenen Zahlen können allerdings nur als grober Anhaltspunkt verstanden werden, weil die Ergebnisse jeweils auf unterschiedlichen Daten erzielt wurden.

Die in dieser Arbeit gemessenen Werte bzgl. der Segmentierungsgenauigkeit sind vergleichbar mit den in der Literatur berichteten Werten, obwohl Verfahren aus der Literatur im direkten Vergleich auf dem in dieser Arbeit verwendeten Daten deutlich schlechter abscheiden. Ein Grund dafür könnte sein, dass der in dieser Arbeit verwendete Evaluationsdatensatz schwierigere Fälle enthält. Es wurden explizit Lymphknoten mit nekrotischem Gewebe, geringem Kontrast und starkem Bildrauschen einbezogen. Aufgrund dieser möglichen Ungleichheit wird auf eine weitere Bewertung verzichtet.

### 5.1.5. Zusammenfassung

Die in dieser Arbeit vorgestellten Verfahren zur Lymphknotensegmentierung wurden in Bezug auf Segmentierungsgenauigkeit, Robustheit und Laufzeit untersucht. Auch ein Vergleich zu in der Literatur vorgestellten Verfahren erfolgte. Grundlage für diese Evaluation war ein aus 49 Lymphknoten bestehender Datensatz, welcher auch stark verrauschte Bilder, Lymphknoten mit zentralem nekrotischen Gewebe und geringem Kontrast an der Lymphknotengrenze enthält.

Die vorgestellten Erweiterungen des Masse-Feder-Modells waren zwar in der Lage, den größten Nachteil des modellbasierten Verfahrens, nämlich die mangelnde Unterstützung unterschiedlicher Lymphknotengrößen, zu verringern, dennoch ist die Segmentierungsgenauigkeit 30% geringer als bei manueller Segmentierung. Das vorgestellte Radialstrahl-basierte 2D-Verfahren zur Segmentierung von Lymphknoten in CT-Schichten zeigte, das Radialstrahl-basierte Verfahren durch die Verwendung von lokalem Formwissen viele unterschiedliche Formen und durch die Verwendung globaler Optimierung unterschiedliche Größen unterstützen.

Bei dem Radialstrahl-basierten 3D-Verfahren wurden folgende Ergebnisse erzielt:

- **Formbeschränkung:** Die prinzipielle Beschränkung auf sternkonvexe Formen ist in der Praxis vernachlässigbar.
- **Vergleich von Optimierungsverfahren:** Bei der Bestimmung der Radien für die einzelnen Strahlen erzielt das Graph-Cut basierte Verfahren die beste durchschnittliche Segmentierungs-

genauigkeit, dicht gefolgt von Belief-Propagation ( $-3\%$ ), aber mit großem Abstand zum Gradientenabstiegsverfahren ( $-26\%$ ).

- **Segmentierungsgenauigkeit:** Im Durchschnitt ist die erzielte Segmentierungsgenauigkeit nur geringfügig schlechter als bei manueller Segmentierung ( $-2\%$ ). Lediglich bei einigen Lymphknoten, die einen sehr geringen Bildkontrast an der Lymphknotengrenze aufweisen oder stark verrauscht sind, werden unzufriedenstellende Segmentierungsergebnisse erzielt. Größe und Form haben keinen signifikanten Einfluss.
- **Vergleich zur Literatur:** Im Vergleich zu existierenden Verfahren erreicht das vorgestellte Verfahren eine deutlich höhere Segmentierungsgenauigkeit (Region-Growing  $-93\%$ , Masse-Feder-Modell  $-29\%$ , Barbu[67]  $-41\%$ , Wang[93]  $-25\%$ ).
- **Robustheit:** Durch die vorgestellte saarpunktinvariante Erweiterung hat die typische Variation des manuell vorgegebenen Saarpunkts einen sehr geringen Einfluss auf das Segmentierungsergebnis ( $< \pm 0.7\%$  in dem Median-Fall).
- **Volumenbestimmung:** Die Volumenbestimmung von Lymphknoten mit dem vorgestellten Verfahren erreicht durchschnittlich eine deutlich geringere Inter-Observer-Variabilität als manuelle Segmentierung ( $-63\%$ ) bei nur geringfügig schlechterer Segmentierungsgenauigkeit ( $-6\%$ ).
- **Bildgradientenbasiertes Formwissen:** Die Verwendung des bildgradientenbasierten Formwissens konnte die durchschnittliche Segmentierungsgenauigkeit des auf Belief-Propagation basierenden Verfahrens etwas verbessern ( $+1\%$ ). Dabei wird ein größerer Segmentierungsgenauigkeitsgewinn üblicherweise, wie erwartet, bei weniger runden Lymphknoten erzielt.
- **Interaktives Verfahren:** Die interaktive Korrektur verbessert das schlechteste Segmentierungsergebnis um  $107\%$  gegenüber dem semi-automatischen Verfahren. Im Durchschnitt erzielt das interaktive Verfahren leicht bessere Ergebnisse als manuelle Segmentierung ( $+3\%$ ). Dazu sind pro Lymphknoten durchschnittlich 1.31 Eingriffe nötig und eine Interaktionsdauer von 12.1 Sekunden muss aufgewendet werden.
- **Laufzeit:** Auf aktueller Hardware benötigt die Segmentierung eines Lymphknotens durchschnittlich 0.1, 2.7 bzw. 3.6 Sekunden unter Verwendung des Gradientenabstiegs-, Belief-Propagation, bzw. des Graph-Cut-Verfahrens.

Die Ergebnisse zeigen, dass das vorgestellte Verfahren einen hohen Reifegrad besitzt und daher in der Praxis erprobt und schlußendlich auch eingesetzt werden kann.

## 5.2. Registrierung von Lymphknoten in Kopf-Hals-Tomographien

### 5.2.1. Material

Grundlage für die Evaluation sind insgesamt 137 Kopf-Hals-Tomographien von 50 Patienten mit Kopf-Hals-Karzinom, welche im klinischen Alltag aufgenommen wurden. Der Aufnahmebereich umfasst hauptsächlich den Hals, in vielen Fällen aber auch noch den Thorax und/oder den Kopf. Die Bildmodalitäten umfassen (kontrastverstärkte) CT- und MRT-Aufnahmen.

Die CT-Aufnahmen haben einen Schichtabstand zwischen 0.5 und 3 *mm* und weisen zum Teil beträchtliche Bildartefakte auf, welche hauptsächlich durch metallhaltige Zahnfüllungen, aber auch durch andere Implantate zustande gekommen sind. Bei den MRT-Aufnahmen handelt es sich morphologische T1 bzw. T2-gewichtete Aufnahmen, welche sowohl in sagittalen, koronalen und axialen Schichten üblicherweise mit einem Schichtabstand zwischen 3 und 5 *mm* aufgenommen wurden. Die Volumen beschränken sich zum großen Teil auf den Mundbereich und weisen oftmals typische MRT-Artefakte, wie z.B. Bewegungsartefakte oder lokal niedrige Signalstärke auf. Einzig die sagittalen Schichten decken einen größeren Teil des Halsbereichs ab und haben daher in der folgenden Evaluierung eine herausragende Bedeutung.

Neben Multimodalität, in diesem Fall also CT- und MRT-Aufnahmen desselben Patienten zum gleichen Zeitpunkt, sind auch Zeitserien vorhanden. Wohingegen bei Ersteren die Annahme getroffen werden kann, dass keine Unterschiede in der Anatomie, sondern nur in der Position des Patienten vorhanden sind, ist dies bei Letzteren nicht möglich. Die erste Aufnahme der Zeitserie wurde in der Regel vor dem chirurgischen Eingriff, eine zweite 6 Monate danach und ggf. eine dritte weitere 6 Monate später aufgenommen. Zwischen der ersten und zweiten Aufnahme sind daher starke Änderungen in der Anatomie aufgrund der Operation zu erwarten. Der Tumor ist entfernt, in vielen Fällen auch Lymphknoten im Halsbereich und Teilbereiche mussten durch plastische Chirurgie rekonstruiert werden.

Alle Aufnahmen wurden ohne irgendwelche Vorrichtungen zur Fixierung des Patienten aufgenommen. Daher variiert die Position der Wirbelsäule von Aufnahme zu Aufnahme und nicht-rigide Registrierungsverfahren werden benötigt, um diese Deformationen auszugleichen.

### 5.2.2. Multimodale rigide Registrierung des Kopfes

Die in Abschnitt 4.2.1 beschriebene initiale vollautomatische rigide Registrierung des Kopfes wurde anhand von insgesamt 81 Aufnahmen von 15 Patienten evaluiert. Diese gliedern sich in 18 CT- und 63 MRT-Aufnahmen (davon 29 mit axialen, 23 mit koronalen und 11 mit sagittalen Schichten). Die Mundhöhle ist in allen Aufnahmen zu sehen, ansonsten sind insbesondere die MRT-Aufnahmen stark abgeschnitten.

**Region of Interest Detektion und Initialisierung:** In allen 81 Aufnahmen wurden ein Punkt auf dem Kinn, sowie obere und untere Grenzen für den Übergang zwischen Kopf- und Halsbereich manuell annotiert (siehe Abbildung 4.21(b)). Die obere Grenze wurde dabei als die höchste Schicht kaudal (unterhalb) des Kinns definiert, die untere Grenze als die höchste Schicht, in welcher die

|           | perfekte<br>Überlappung | geringfügige<br>Ungenauigkeiten | Misserfolg | durchschnittliche<br>Registrierungsdauer |
|-----------|-------------------------|---------------------------------|------------|--|
| Methode A | 51                      | 1                               | 5          | 11.3 s                                   |
| Methode B | 46                      | 1                               | 10         | 11.7 s                                   |
| Methode C | 45                      | 3                               | 9          | 12.1 s                                   |
| Methode D | 6                       | 7                               | 44         | 6.0 s                                    |

Tabelle 5.13.: Vergleich der Registrierungsqualität unterschiedlicher rigider Registrierungsverfahren für 57 CT/MRT-Bildpaare

Ansätze der Schulter zu erkennen sind. Durchschnittlich umfasste der so definierte Bereich 24.4 mm.

Die Heuristik zur automatischen Detektion des Kopf-Hals-Übergangs und der Kinnschuppe wurde dann auf alle 81 Aufnahmen angewendet. Anschließend wurde sowohl der Abstand des detektierten Übergangs zu dem zuvor definierten gewünschten Bereich als auch der Abstand von der detektierten und manuell annotierten Kinnschuppe gemessen.

In 56 Fällen wird der Übergangsbereich getroffen, in den anderen Fällen beträgt der durchschnittliche Abstand 17.5 mm. Der durchschnittliche Fehler der Kinnschuppendetektion beträgt 9.6 mm, wobei in 76 Fällen der Abstand weniger als 20 mm beträgt. In den restlichen 5 Aufnahmen beträgt dieser im schlechtesten Fall 74.8 mm.

Die Begutachtung dieser Fälle zeigt, dass andere Merkmale des Kopfes, wie z.B. die Lippe, fälschlicherweise detektiert wurden. Die Gründe dafür waren in allen 5 Aufnahmen ausnahmslos starke Bildartefakte, welche zu einer falschen Detektion des Kopf-Hals-Übergangs führten. Auch wenn diese Fehler zunächst hoch erscheinen mögen, ist die Initialisierung der Registrierung mit dieser Landmarke nach wie vor robuster als eine geometrie- oder momentenbasierende Initialisierung, weil Letztere stark beschnittene Aufnahmen nicht unterstützt.

**Bildbasierte Registrierung:** Rigide Registrierung kann durch den Vergleich der durch die Registrierung erhaltenen Transformation mit einer Referenztransformation durchgeführt werden [151]. Die Referenztransformation wird dabei aus korrespondierenden Landmarken berechnet. Alternativ kann die Korrespondenz von knöchernen Strukturen herangezogen werden [152]. Zur Evaluation des in dieser Arbeit entwickelten rigiden Registrierungsverfahrens sind ungeeignet. Ersteres, weil bei klinisch akquirierten Daten keine Referenztransformation zur Verfügung steht und Letzteres, weil sich ein Teil der knöchernen Strukturen in einem potentiell deformierten Bereich (z.B. Wirbelsäule) befindet.

Daher wurde stattdessen auf eine visuelle Begutachtung zurückgegriffen. Aus den insgesamt 81 verfügbaren Aufnahmen können 57 zueinander passende Paare von Aufnahmen desselben Patienten gebildet werden. Alle Paare wurden mit der kompletten in Abschnitt 4.2.1 beschriebenen Methode (*Methode A*) registriert. Um die Notwendigkeit der einzelnen Teilschritte des Verfahrens zu zeigen, wurden in weiteren Experimenten unterschiedliche Schritte weggelassen (*Methode B*: Keine Vorverarbeitung; *Methode C*: Keine Vorverarbeitung und keine Rotationsbestimmung vor der vollrigiden Registrierung). Zum Vergleich wurden alle Paare im jeweiligen Bildschwerpunkt initialisiert und durch einfache Maximierung der Mutual Information der kompletten Aufnahmen rigide registriert (*Methode D*). Um Voreingenommenheit des Beobachters (erfahrener Bildverarbeiter) auszuschließen, wurden

|                     | einfacher Gradientenabstieg |         | adaptiver Gradientenabstieg |         |
|---------------------|-----------------------------|---------|-----------------------------|---------|
|                     | absolut                     | relativ | absolut                     | relativ |
| korrekte Anpassung  | 41                          | 71%     | 50                          | 86%     |
| verrutschter Wirbel | 14                          | 24%     | 7                           | 12%     |
| Fehlschlag          | 3                           | 5%      | 1                           | 2%      |
| Gesamt              | 58                          | 100%    | 58                          | 100%    |

Tabelle 5.14.: Visuelle Beurteilung der Anpassungsqualität des artikulierte Atlas

die Registrierungsergebnisse in zufälliger Reihenfolge angezeigt, ohne dass er wusste, mit welcher der 4 Methoden das Ergebnis erzielt wurde. Der Beobachter sollte dann die Qualität des Registrierungsergebnisses im Kopf in mindestens 2 axialen Schichten beurteilen und in folgende Kategorien einteilen: *perfekte Überlappung*, *geringfügige Ungenauigkeiten* und *Misserfolg*. Deformierbare Bereiche im Hals sollen dabei nicht berücksichtigt werden.

Die Ergebnisse sind in Tabelle 5.13 dargestellt. In 12 Fällen schlug die Detektion des Kinns in einer der beiden Aufnahmen fehl, was zu einer schlechten Initialisierung führt. In 5 dieser Fälle konnte dies allerdings nicht durch die darauf folgende bildbasierte Registrierung ausgeglichen werden, was zeigt, dass eine gute Initialisierung der Transformation für den Erfolg der Registrierung entscheidend ist. Wie man sieht, schlägt die Registrierung ohne Vorverarbeitung (*Methode B*) in einer größeren Anzahl an Fällen fehl (10), was dafür spricht, dass die Beschränkung der Grauwerte auf das Weichteilgewebefenster den Konvergenzbereich des Verfahrens erhöht. Der Unterschied zwischen dem Verfahren ohne Rotationsoptimierung zum zweistufigen Optimierungsverfahren (*Methode B* vs. *Methode C*) fällt nicht so stark aus, wie während der Entwicklung erwartet wurde. Dies kann an der Ungenauigkeit der Detektion des Kinns liegen, was dazu führt, dass eine 3D-Rotation nicht ausreichend ist, um ein gutes Registrierungsergebnis zu erzielen. Trotzdem kann durch das zweistufige Verfahren die Registrierungsgeschwindigkeit erhöht werden. Der naive Ansatz (*Methode D*) erzielt sehr schlechte Ergebnisse bei stark abgeschnittenen Aufnahmen, die in Teilbereichen deformiert sind und somit nicht der Annahme eines starren Körpers gerecht werden.

### 5.2.3. Artikulierter Atlas zur Segmentierung einzelner Knochen

**Anpassungsgüte:** Die Evaluierung der Anpassungsgüte des artikulierte Atlas erfolgte anhand von 58 nicht zum Training des artikulierte Atlas verwendeten CT-Datensätzen von 43 Patienten mit Kopf-Hals-Karzinom. Einige der Aufnahmen weisen eine pathologische Anatomie auf, die entweder durch einen chirurgischen Eingriff, den Tumor oder durch Osteoporose verursacht wurde.

Zunächst wurde der aus 15 (30 inklusive der gespiegelten Bilder) Trainingsdatensätzen erstellte artikulierte Atlas an diese Datensätze mit denselben Parametern angepasst. Die Parameter wurden dabei in informellen Experimenten festgelegt. Die zwei in Abschnitt 4.2.3 vorgestellten Anpassungsstrategien wurden miteinander verglichen: a) einfacher Gradientenabstieg und b) adaptiver Gradientenabstieg. Anschließend bestimmte ein Beobachter (erfahrener Bildverarbeiter) für jede Anpassungsstrategie inwieweit die einzelnen Atlaselemente sich an die jeweilige Struktur korrekt angepasst haben. Konkret wurde für jeden Testdatensatz bestimmt, ob a) alle Elemente korrekt angepasst wurden, b) einige Elemente verrutscht sind (bspw. dass das Element C4 sich an den ähnlich aussehenden Wirbel C5 anpasst) oder c) der Atlas sich an eine komplett andere Bildstruktur angepasst hat. Eine quantitative

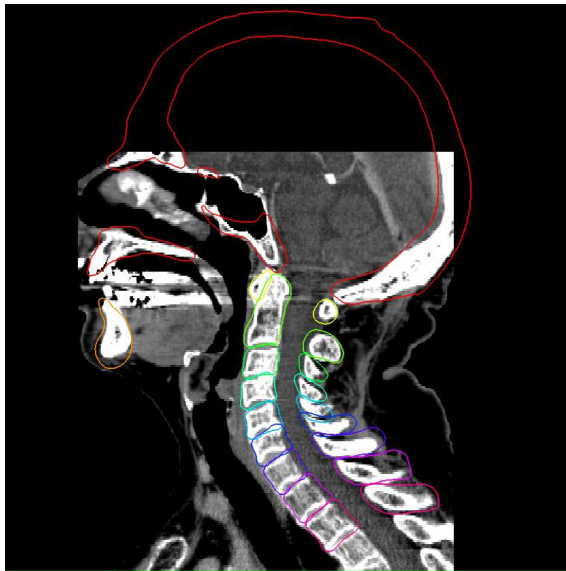
Bestimmung der Segmentierungsgenauigkeit für jeden Knochen erfolgte zunächst nicht, weil i) keine Referenzsegmentierung für diese große Anzahl an Datensätzen vorliegt und ii) die genaue Abgrenzung der einzelnen Knochen im Anwendungsfall dieser Arbeit – multi-rigide Registrierung – im Vergleich zur groben globalen Lokalisation nur eine untergeordnete Rolle spielt.

Wie aus der Tabelle 5.14 hervorgeht, hat der artikulierte Atlas auf klinischen CT-Datensätzen eine Erfolgsrate von ca. 86% unter Verwendung des adaptiven Gradientenabstiegs. Dieser Wert hat sich gegenüber dem einfachen Gradientenabstieg bei sonst konstanten Parametern um fast 25% verbessert. Der gewünschte Effekt des adaptiven Gradientenabstiegs wurde damit also bestätigt. Trotz allem ist die Hauptursache für eine nicht vollständig korrekte Anpassung nach wie vor das Konvergieren mindestens eines Wirbelelements in einen falschen benachbarten und sehr ähnlich aussehenden Wirbel. Dies tritt insbesondere bei unüblicher Anatomie (siehe Abbildung 5.15(c)), wie z.B. dem Fehlen einer Bandscheibe oder aber bei unüblichen Wirbelsäulenpositionen auf, welche nur unzureichend durch den gelernten Artikulationsraum abgedeckt wurden (siehe Abbildung 5.15(d)). Eine Erhöhung der Anzahl an Trainingsdatensätzen kann bei Letzterem behilflich sein.

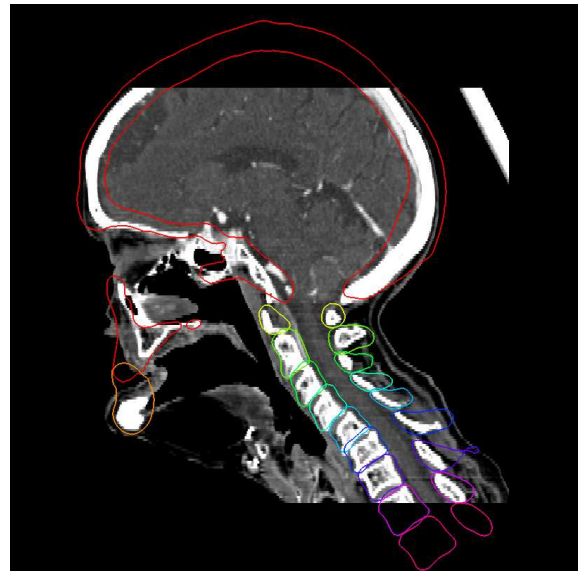
In einem Fall schlägt die Anpassung komplett fehl. Dies ist der unzureichenden Initialisierung geschuldet, was dazu führt, dass das lokale Anpassungsverfahren in einem falschen lokalen Minimum des Energiefunktionals aus Gleichung 4.68 konvergiert. Eine bessere Initialisierung, bspw. durch ein globales Wirbeldetektionsverfahren (z.B. [153]), könnte dieses Problem beheben. In nur einem Fall passiert es, dass das Ergebnis des einfachen Gradientenabstiegs dem des adaptiven Gradientenabstiegs überlegen ist. In diesem Fall waren die Wirbel C3, C4 und C5 stark verkleinert und ohne Bandscheibe sehr nahe beieinander. Dies führte dazu, dass die Atlaselemente C3 und C4 alle drei Wirbel abdeckten und damit als direkte Folge des adaptiven Vorgehens alle folgenden Wirbel verrutscht waren. Bei dem einfachen Gradientenabstieg hingegen befanden sich die Atlaselemente Th1 und Th2 zufällig sehr nahe an den korrekten Wirbeln, was dazu führte, dass die sich weiter oben befindlichen Elemente nicht verrutschen konnten.

**Segmentierungsgenauigkeit:** In einem zweiten Experiment wurde nun die Segmentierungsgenauigkeit quantitativ durch Leave-One-Out Kreuzvalidierung bestimmt. Dazu wurden insgesamt 15 verschiedene artikulierte Atlasse durch das in Abschnitt 4.2.2 beschriebene Verfahren erstellt, wobei jeweils nur 14 der insgesamt 15 vorliegenden Datensätze mit manuellen Annotationen zum Training verwendet wurden. Dieser Atlas wurde dann an den jeweils ausgelassenen, also nicht zum Training verwendeten Datensatz, angepasst. Anschließend wurden die Segmentierungsergebnisse der einzelnen Elemente mit den vorliegenden manuellen Segmentierungen verglichen. Als Ähnlichkeits-/Fehlermaße kamen dabei die schon zur Lymphknotensegmentierung verwendeten und in Abschnitt 5.1.1 beschriebenen volumetrischen Maße *Dice Similarity Coefficient* (DSC) und *Volumetric Overlap Error* (VOE) sowie die abstandsbasierenden Maße *Average Surface Distance* (ASD) und die symmetrische *Maximum Surface Distance* (MSD) zum Einsatz.

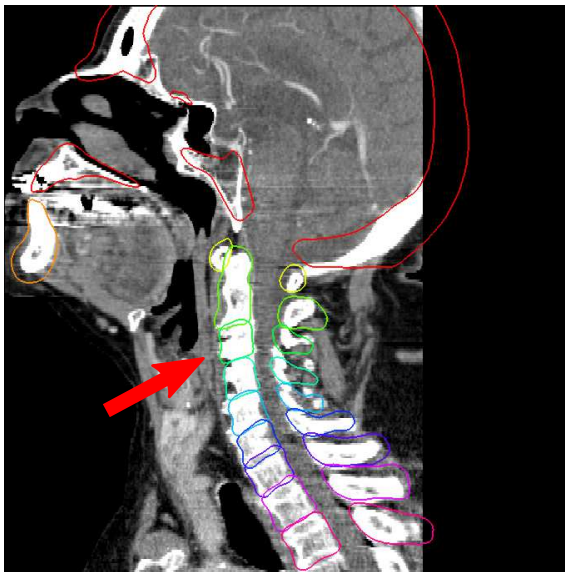
Tabelle 5.15 zeigt die Ergebnisse. Auch hier lag die Erfolgsrate bei 86%. Da die nicht korrekt angepassten Elemente nicht aussagekräftig für die Segmentierungsgenauigkeit sind, wurden diese einmal mit und einmal ohne Einbeziehung dieser Elemente berechnet. Für die korrekt angepassten Elemente beträgt der durchschnittliche ASD für die Wirbel 0.65 mm. Im Vergleich zu einem Formmodell-basierten Ansatz [145] ist dies nur geringfügig schlechter. Bei der MSD werden sogar leicht bessere Ergebnisse erzielt (6.51 vs. 6.88 mm). Auffällig ist der hohe MSD des Schädels und des Unterkiefers. Der Grund dafür liegt zum Teil an starken Metallartefakten in den Zähnen der Patienten. Diese wurden



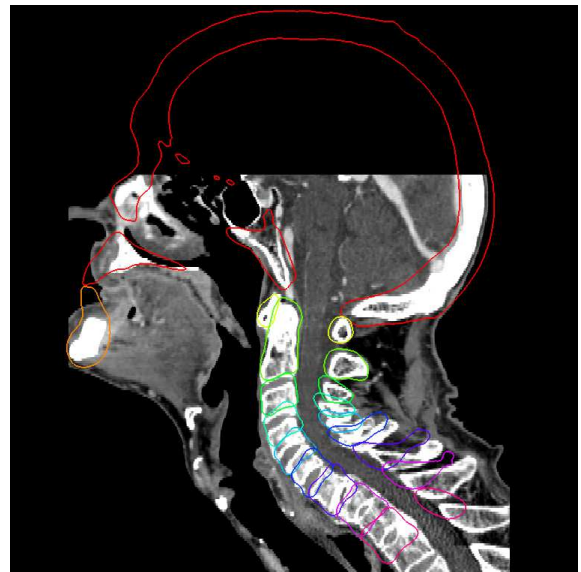
(a) Korrekte Anpassung



(b) Korrekte Anpassung



(c) Verrutschte Wirbel aufgrund verbundener Wirbel (C3 und C4)



(d) Verrutschte Wirbel aufgrund stark gekrümmter Wirbelsäule, die nicht vom Artikulationsraum erfasst wurde

Abbildung 5.15.: Beispiele für den angepassten artikulierten Atlas. Die Anpassung erfolgte durch das adaptive Gradientenabstiegsverfahren.

|             | Erfolgs-<br>rate | Alle Fälle |       |      |       | Erfolgreiche Anpassung |       |      |       |
|-------------|------------------|------------|-------|------|-------|------------------------|-------|------|-------|
|             |                  | DSC        | VOE   | ASD  | MSD   | DSC                    | VOE   | ASD  | MSD   |
| Schädel     | 100%             | 0.87       | 22.76 | 0.68 | 16.15 | 0.87                   | 22.76 | 0.68 | 16.15 |
| Unterkiefer | 100%             | 0.85       | 26.54 | 0.93 | 10.47 | 0.85                   | 26.54 | 0.93 | 10.47 |
| Atlas       | 100%             | 0.82       | 30.38 | 0.68 | 5.95  | 0.82                   | 30.38 | 0.68 | 5.97  |
| Axis        | 93%              | 0.84       | 26.39 | 0.72 | 6.14  | 0.86                   | 23.84 | 0.61 | 5.66  |
| C3          | 93%              | 0.77       | 34.75 | 1.00 | 6.61  | 0.82                   | 30.33 | 0.67 | 5.95  |
| C4          | 80%              | 0.75       | 36.13 | 1.19 | 7.18  | 0.84                   | 27.21 | 0.56 | 5.42  |
| C5          | 80%              | 0.72       | 38.60 | 1.63 | 8.65  | 0.84                   | 27.07 | 0.57 | 6.24  |
| C6          | 80%              | 0.69       | 41.33 | 2.01 | 10.20 | 0.83                   | 28.44 | 0.65 | 7.29  |
| C7          | 73%              | 0.67       | 42.58 | 2.41 | 12.13 | 0.85                   | 26.58 | 0.66 | 8.79  |
| Th1         | 73%              | 0.66       | 43.68 | 2.94 | 11.78 | 0.85                   | 26.04 | 0.72 | 6.80  |
| Th2         | 73%              | 0.64       | 45.51 | 3.38 | 12.21 | 0.84                   | 27.70 | 0.79 | 6.47  |
| Wirbel      | 83%              | 0.73       | 37.71 | 1.77 | 8.99  | 0.84                   | 27.51 | 0.65 | 6.51  |
| Gesamt      | 86%              | 0.75       | 35.33 | 1.60 | 9.77  | 0.84                   | 26.99 | 0.68 | 7.75  |

Tabelle 5.15.: Segmentierungsqualität des artikulierten Atlas

teilweise fälschlicherweise aufgrund der hohen Bildintensität als Schädel- oder Unterkieferknochen klassifiziert. Auch kam es beim Unterkiefer teilweise zu starker Untersegmentierung, weil Zähne in den Trainingsdaten nur sehr selten vorhanden waren und so diese bei der Segmentierung zum Teil nicht berücksichtigt werden konnten. Die Annahme, dass die Formvarianz eines Atlaselements gering genug ist, dass sie durch ein einfaches Wahrscheinlichkeitsbild repräsentiert werden kann, ist hier also nur unzureichend erfüllt. Für das Anwendungsszenario dieser Arbeit – der multi-rigiden Registrierung – spielt die Genauigkeit jedoch nur eine untergeordnete Rolle. Viel wichtiger ist die Erfolgsrate (siehe Tabelle 5.14).

#### 5.2.4. Registrierung einzelner Knochen

Im nächsten Experiment soll die Registrierungsgenauigkeit der einzelnen zuvor durch den artikulierten Atlas bestimmten rigid Elementen gemessen werden. Datengrundlage dazu sind 20 MRT-Aufnahmen und 16 CT-Zeitserienaufnahmen von Patienten, für die bereits eine erfolgreiche Segmentierung durch den artikulierten Atlas vorliegt.

Bei den MRT-Aufnahmen handelt es sich um T1-gewichtete Aufnahmen in sagittalen Schichten mit einem Schichtabstand von 3 – 5 mm, welche den Kopf und die Halswirbelsäule darstellen. Da die Aufnahmen in etwa zum selben Zeitpunkt wie die dazugehörigen CT-Aufnahmen aufgenommen wurden, kann davon ausgegangen werden, dass die Anatomie innerhalb der CT/MRT-Paare keine Änderungen aufweist. Sämtliche Unterschiede in den Aufnahmen resultieren also aus etwaigen Unterschieden in der Position des Patienten.

Die CT-Zeitserienaufnahmen besitzen, wie die original CT-Aufnahmen auch, einen Schichtabstand von 1 – 2 mm. Diese Aufnahmen wurden allerdings zu anderen Zeitpunkten als die original CT-Aufnahmen aufgenommen und zwar in 9 Fällen nach 6 Monaten, in 5 Fällen nach 12 Monaten und in weiteren 3 Fällen nach 18 Monaten. Bei all diesen Patienten wurde kurz nach der Originalaufnahme ein chirurgischer Eingriff vorgenommen. Insbesondere dadurch, aber auch durch den teilweise



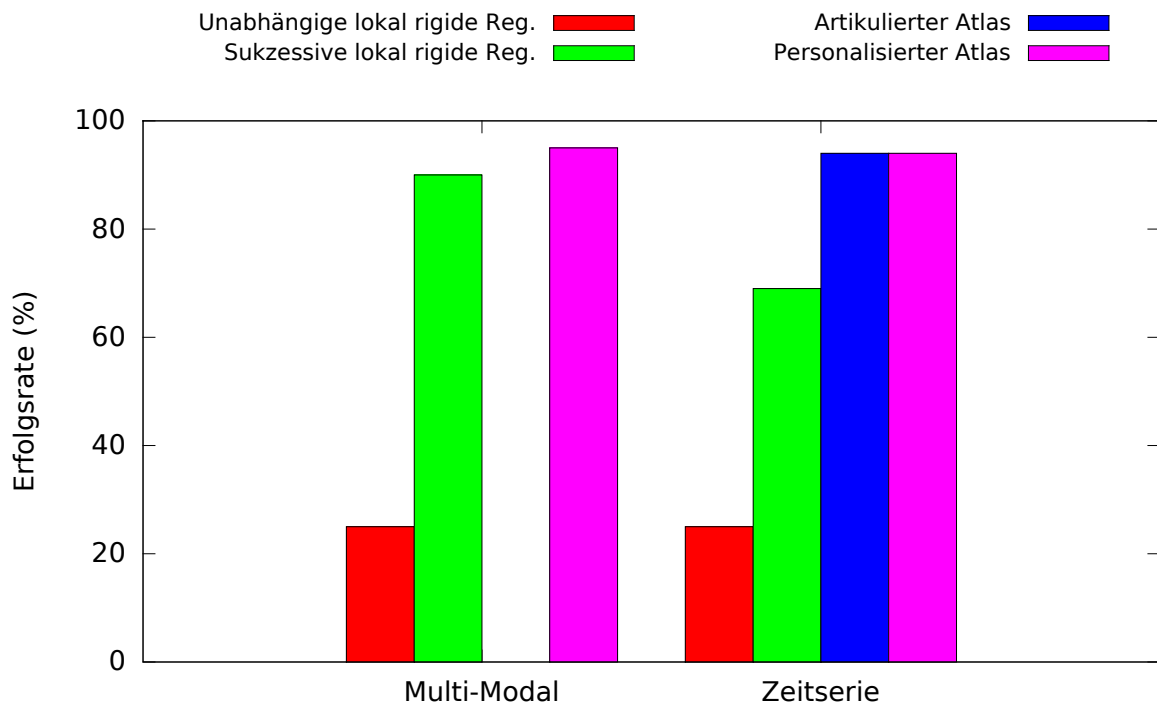


Abbildung 5.16.: Erfolgsrate der Registrierung einzelner Knochen durch verschiedene Verfahren

langen Abstand zwischen den Aufnahmen, muss von einer signifikanten Änderung in der Anatomie ausgegangen werden.

Zur Initialisierung wurde der Kopf in den MRT- und den CT-Zeitserienaufnahmen zunächst rigide mit der in Abschnitt 4.2.1 beschriebenen und in Abschnitt 5.2.2 evaluierten Methode registriert. Anschließend wurde durch folgende Verfahren für jedes Atlaselement je eine rigide Transformation gefunden:

- **Unabhängige lokal rigide Registrierung:** Jedes Element wird separat von allen durch die Maximierung von Mutual Information registriert. Die Initialisierung der Transformationen erfolgt für alle Elemente durch die globale rigide Registrierung.
- **Sukzessive lokal rigide Registrierung:** Die Elemente werden nacheinander beginnend mit Schädel bis hin zum Wirbel Th2 lokal rigide durch Maximierung von Mutual Information registriert, wobei die Initialisierung der Transformation jeweils durch das Registrierungsergebnis des vorherigen Elements erfolgt (Siehe Abschnitt 4.2.4). Wissen über die Artikulation wird nicht eingebracht.
- **Artikulierter Atlas:** Der artikulierte Atlas wird durch das adaptive Gradientenabstiegsverfahren an die Aufnahme unter Berücksichtigung des gelernten Artikulationsraums angepasst. Aufgrund der unterschiedlichen Artikulationen in den Aufnahmen werden dann rigide Transformationen der einzelnen Elemente berechnet. Da der artikulierte Atlas nur auf CT-Aufnahmen angewendet werden kann, kann dieses Verfahren nur zur Registrierung von CT-Zeitserienaufnahmen, nicht jedoch zur Registrierung der MRT-Aufnahmen eingesetzt werden.

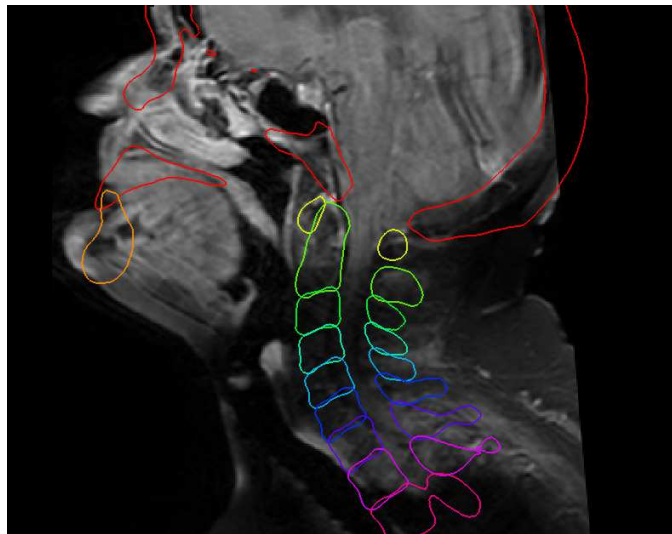


Abbildung 5.17.: Aufgrund von fehlender Signalstärke in der MRT-Aufnahmen konnten die Knochen der Wirbelsäule in diesem Beispiel mit keinem Verfahren (hier der personalisierte artikulierte Atlas) korrekt registriert werden.

- **Personalisierter Atlas:** Ein durch die originalen CT-Aufnahme personalisierter Atlas wird durch das adaptive Gradientenabstiegsverfahren an die MRT- bzw. CT-Zeitserienaufnahme angepasst.

Für die einzelnen Registrierungsergebnisse wurde dann von einem Beobachter (erfahrener Bildverarbeiter) beurteilt, ob sich das jeweilige Element aus der original CT-Aufnahme an der richtigen Stelle mit der richtigen Orientierung in der MRT- bzw. CT-Zeitserienaufnahme befindet.

Die Erfolgsrate ist in Abbildung 5.16 dargestellt. Es zeigt sich, dass durch globales Wissen über die Lage der rigiden Strukturen zueinander mit Hilfe des (personalisierten) artikulierten Atlas sowohl bei multimodaler Registrierung als auch bei der Registrierung von Zeitserien die besten Ergebnisse erzielt werden können. Im Falle der CT/MRT-Registrierung schlug dies in nur einem der 20 Fälle fehl (siehe Abbildung 5.17). Der Grund dafür war in diesem Fall eine extrem niedrige Signalstärke der MRT-Aufnahme im Bereich der Wirbelsäule. Im CT/MRT-Fall schnitt das sukzessiv lokal rigide Verfahren fast so gut wie der personalisierte artikulierte Atlas ab. Im Zeitserienfall hingegen schnitt es mit nur 69%-iger Erfolgsrate deutlich schlechter ab. Dort wurde der Unterkiefer oft nicht korrekt registriert. Der Grund dafür waren in der Regel starke Änderungen in der Anatomie des Unterkiefers durch den chirurgischen Eingriff. Wie zu erwarten, wurden die schlechtesten Ergebnisse von dem unabhängigen lokal rigiden Verfahren erzielt. Dieses war bei beiden Bildtypen nur in 25% der Fälle in der Lage, eine korrekte Registrierung für alle Strukturen liefern. Dies war immer dann der Fall, wenn die Deformationen zwischen den Aufnahmen sehr gering waren und sich so die Initialisierung bereits nah genug an der gesuchten Transformation befand. Ansonsten konnten insbesondere diejenigen Wirbel, die sich weit entfernt von dem initial rigide registrierten Kopf befanden, nicht korrekt registriert werden. Dennoch wird diese Art der Registrierung von in der Literatur vorgestellten multi-rigidem Verfahren überwiegend verwendet.

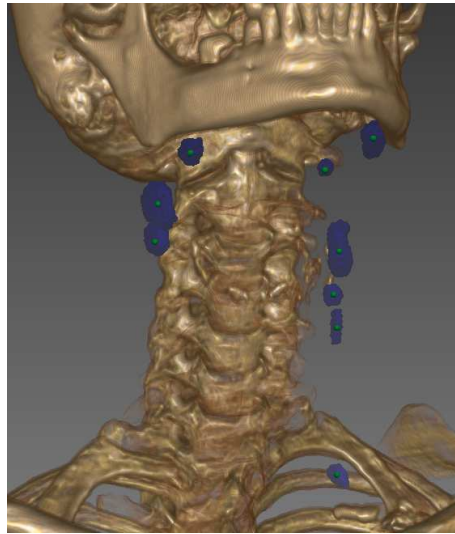


Abbildung 5.18.: Verteilung der Registrierungszielorgane im Weichteilgewebe

### 5.2.5. Weichteilgewebe-Deformation

Sind die einzelnen Knochen erfolgreich registriert, sollen nun die Registrierungsgenauigkeit im Weichteilgewebe, insbesondere in den Lymphknoten und die Qualität der deformierten Aufnahme bestimmt werden.

**Registrierungsgenauigkeit:** Allgemein gilt die Bestimmung der Registrierungsgenauigkeit von deformierbaren Registrierungsverfahren als ein schwieriges Problem, weil auf klinischen Daten kein Referenzdeformationsfeld vorhanden ist, gegen welches ein Fehlermaß bestimmt werden könnte. Ist die Anatomie in beiden Aufnahmen unterschiedlich, also bei Inter-Subjekt-Registrierung, aber auch bei Zeitserienaufnahmen, ist noch nicht einmal die eindeutige Existenz eines solchen Referenzdeformationsfeldes sichergestellt. Künstlich generierte Deformationsfelder, welche erst auf eine Aufnahme angewendet werden, um diese bekannte Deformation dann durch ein Registrierungsverfahren wieder rückgängig zu machen, haben einen anderen Nachteil. Wird dasselbe Modell sowohl zur Deformationsfeldgenerierung, als auch zur Registrierung verwendet, kann das Registrierungsverfahren nicht unabhängig evaluiert werden. Wird hingegen ein anderes Verfahren für beide Prozesse verwendet, kann nicht bestimmt werden, ob ein Fehler bei der Registrierung oder schon bei der Deformationsfeldgenerierung gemacht wurde.

Ist kein Referenzdeformationsfeld vorhanden, wird oft auf das Messen von Hilfsmaßen zurückgegriffen. Dazu wird oft auf Bildähnlichkeit, welche bspw. durch Mutual Information gemessen werden kann, aber auch auf Gewebeüberlappungskoeffizienten zurückgegriffen. Diese Hilfsmaße sind allerdings nur schwach mit dem tatsächlichen Registrierungsfehler korreliert und sollten daher nicht als Gütekriterium angegeben werden [154].

Die einzige verbleibende Möglichkeit der quantitativen Evaluierung der Registrierungsgenauigkeit besteht im Messen des Registrierungsfehlers an korrespondierenden Landmarken. Dieser Fehler wird als *Target Registration Error* (TRE) bezeichnet. Da ein Ziel dieser Arbeit die Registrierung von

|             | Deformation | Transform. | 2-Stufen | Oberfläche | Rigide | B-Spline |
|-------------|-------------|------------|----------|------------|--------|----------|
| Minimum     | 0.83        | 0.64       | 0.12     | 0.92       | 1.33   | 0.43     |
| 0.25-Quant. | 2.84        | 3.06       | 3.06     | 2.97       | 3.95   | 4.79     |
| Median      | 4.38        | 5.04       | 4.56     | 4.17       | 6.48   | 7.04     |
| 0.75-Quant. | 6.70        | 6.84       | 6.85     | 6.12       | 8.59   | 10.33    |
| Maximum     | 23.07       | 26.80      | 26.27    | 25.37      | 61.21  | 30.05    |
| Ø           | 5.49        | 5.52       | 5.38     | 5.05       | 8.34   | 8.01     |
| Std. Abw.   | 3.93        | 3.70       | 3.64     | 3.55       | 8.16   | 5.51     |

Tabelle 5.16.: Target-Registration-Error von Lymphknotenzentren für unterschiedliche Arten der Deformationsfeldberechnung in mm

Lymphknoten ist, wurden diese als Zielorgan ausgewählt. Weitere Gründe, die für die Auswahl dieser Strukturen als Zielorgan sprechen, sind, dass diese vergleichsweise gleichmäßig im Weichteilgewebe des Kopf-Hals-Bereichs verteilt sind (siehe Abbildung 5.18), lokal beschränkt sind und in beiden Bildmodalitäten sichtbar sind. In den 19 der 20 Paare von CT- und MRT-Aufnahmen, bei denen die Knochen erfolgreich registriert werden konnten, wurden durchschnittlich je 5.2 korrespondierende Lymphknoten identifiziert. Diese wurden in der CT-Aufnahme durch das in Abschnitt 4.1.3 beschriebene Radialstrahl-basierte Verfahren und in der MRT-Aufnahme manuell segmentiert. Anschließend wurde dann nicht die Überlappung der gesamten Segmentierungen bestimmt, sondern lediglich der Abstand des Schwerpunkts der Lymphknoten in der CT- und der deformierten MRT-Aufnahme gemessen. Aufgrund der unterschiedlichen räumlichen Auflösungen der beiden Aufnahmen erwies sich dies als robuster. In den Zeitserien-CT-Aufnahmen kann dieses Fehlermaß jedoch nicht bestimmt werden, da in vielen Fällen gut sichtbare Lymphknoten chirurgisch entfernt wurden oder andere Anatomieänderungen vorhanden sind, welche nicht durch das Registrierungsverfahren ausgeglichen werden sollen.

Der Target-Registration-Error wurde für unterschiedliche Arten der Berechnung des Deformationsfeldes aus den Transformationen der Knochen (vgl. Abschnitt 4.2.6) bestimmt:

- **Propagierung der Deformation:** Für alle Voxel, die innerhalb der Knochen liegen, für welche eine rigide Transformation gefunden wurde, wurde die Deformation berechnet. Aus den bekannten Deformationen an diesen Stellen wird dann ein dichtes Deformationsfeld durch die Lösung der Laplacegleichung in einem Schritt berechnet.
- **Propagierung der Transformationen:** Anstatt der Deformation, werden auch die Rotationsanteile in des Weichteilgewebe durch Lösung der Laplacegleichung propagiert. Wieder erfolgt die Berechnung in einem Schritt.
- **2-stufige Propagierung der Transformationen:** In einem ersten Schritt beschränkt sich die Berechnung des Transformationsfeldes auf den Bereich innerhalb des Körpers. Erst in einem zweiten Schritt wird auch das Deformationsfeld außerhalb des Körpers berechnet.
- **Körperoberfläche:** Nach der 2-stufigen Berechnung des Transformationsfeldes werden hier zusätzlich die aus beiden Aufnahmen extrahierten Körperoberflächen in Deckung gebracht.

Zum Vergleich wurde der Fehler auch für das in Abschnitt 4.2.1 beschriebene rigide Kopfregistrierungsverfahren und ein rein bildbasiertes hierarchisches B-Spline Registrierungsverfahren [57], welches die Mutual Information beider Aufnahmen maximiert, berechnet. Letzteres ist das Standardverfahren

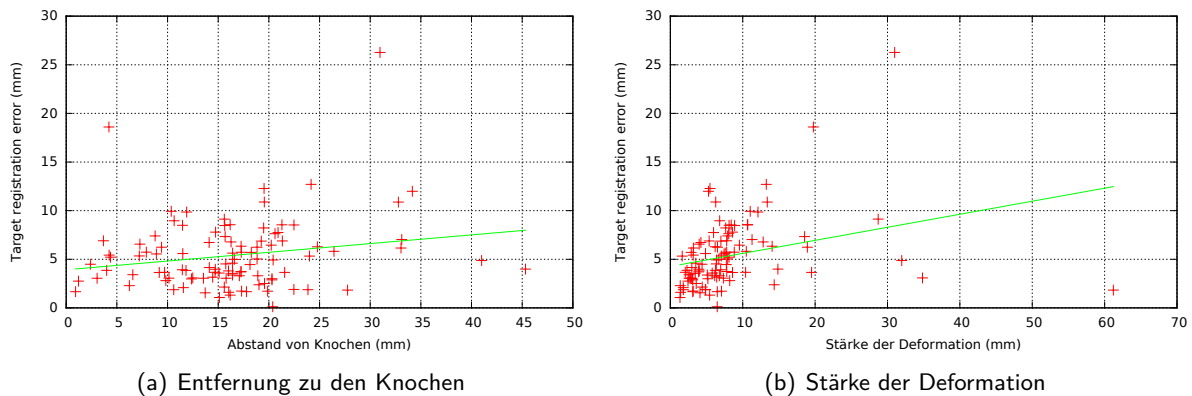


Abbildung 5.19.: Target-Registration-Error der 2-Stufen Anpassung in Abhängigkeit verschiedener Eigenschaften

zur nicht-rigiden multimodalen Registrierung und wurde in vielen Anwendungsgebieten erfolgreich eingesetzt. Hier wurde die Implementierung der *elastix*-Toolbox [155] verwendet.

Der durchschnittliche Target-Registration-Error (TRE) beträgt  $5.49 \text{ mm}$  bei der Propagierung des einfachen Deformationsfeldes,  $5.52 \text{ mm}$  bei der einstufigen Propagierung der rigiden Transformationen,  $5.38 \text{ mm}$  bei der zweistufigen Propagierung der Transformationen und schließlich  $5.05 \text{ mm}$  bei der zusätzlichen Berücksichtigung der Hautoberfläche. Sie schneiden also alle sehr ähnlich ab. Weitere Statistiken über den Target Registration Error sind in Tabelle 5.16 eingetragen. Im Vergleich zur rigiden Registrierung des Kopfes verringert sich der durchschnittliche TRE um 34%. Im Gegensatz dazu verringert das rein bildbasierte Verfahren diesen Fehler nur um 4%. Interessanterweise stellt sich durch die Propagierung der Transformationen im Vergleich zur Propagierung der Deformationen zunächst kein Gewinn ein, erst der zweistufige Propagierungsprozess ermöglicht dies, wenn auch nur geringfügig. Das Einbeziehen der Oberfläche ist in Bezug auf die Registrierungsgenauigkeit sinnvoll.

Im Folgenden soll untersucht werden, wovon der verbleibende Registrierungsfehler abhängt. Dazu wurde in Abbildung 5.19 für die zweistufige Propagierung des Transformationsfeldes der Fehler in Abhängigkeit von zwei Eigenschaften, nämlich der Deformationsstärke eines Targets und der Entfernung des Targets zum nächsten lokal rigide registrierten Knochen, für alle Targets aufgetragen. Die Deformationsstärke an einem bestimmten Target ist dabei durch den euklidischen Abstand zwischen dem Target in dem Referenzbild und demselben Target in dem rigide registrierten Modellbild gegeben. Wie zu sehen ist, nimmt der Fehler tendenziell sowohl mit der Stärke der Deformation als auch mit der Entfernung zu den Knochen zu. Dies war prinzipbedingt zu erwarten, weil bei einer starken Deformation und hoher Entfernung zu den Stützstrukturen die Unzugänglichkeiten des Transformationspropagierungsverfahren stärker zum Tragen kommen. Eine biomechanische Modellierung des Weichteilgewebes könnte zu einer Verbesserung des Ergebnisses führen, erhöht aber die Komplexität des Verfahrens ungemein und verlangt gleichzeitig ein umfangreiches Vorwissen über das Weichteilgewebe, was unter Umständen zunächst einer Segmentierung einer großen Anzahl relevanter Strukturen bedarf.

Der maximale Fehler des Verfahrens ist mit  $26.27 \text{ mm}$  nach wie vor sehr groß. Dieser Wert wurde für einen Lymphknoten in der Nähe des Schlüsselbeins gemessen. In diesem Fall hatten die Schultern

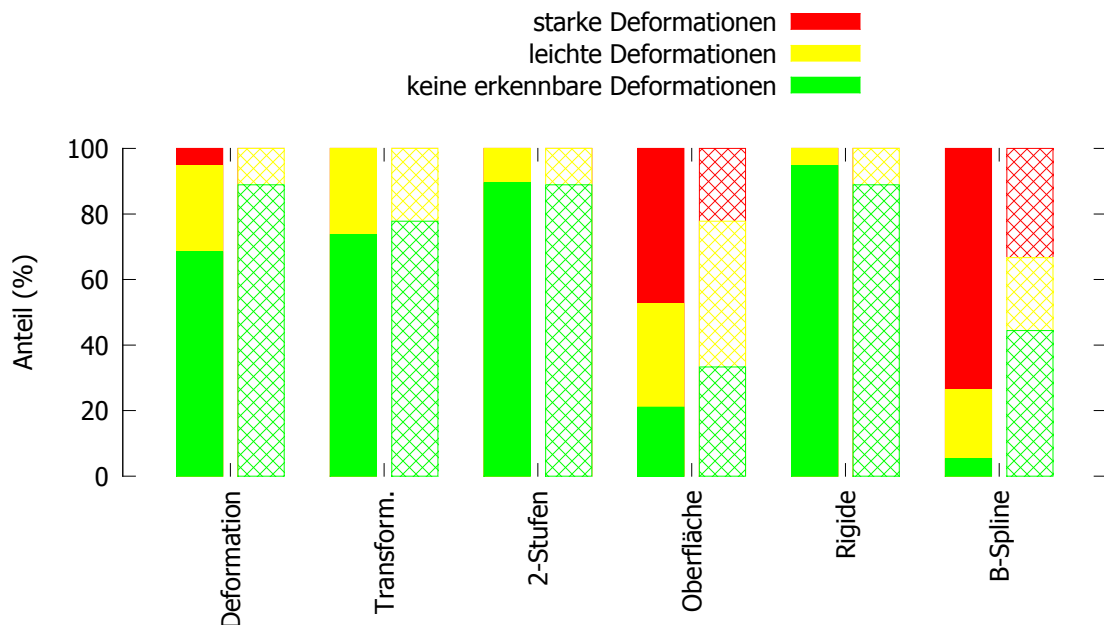


Abbildung 5.20.: Beurteilung der visuellen Qualität der deformierten Aufnahmen für verschiedene Arten der Berechnung des Deformationsfeldes im multimodalen (gefüllt) und Zeitserien-Fall (Muster)

in beiden Aufnahmen eine komplett andere Position. Da die Transformationen jedoch lediglich auf Grundlage der Knochen der Halswirbelsäule und des Kopfes berechnet wurden, konnte die Schulterposition nicht korrigiert werden. Eine Einbeziehung weiterer Knochen, wie z.B. der obersten Rippen oder aber der Knochen im Schulterbereich, könnte diese Art von Fehler verringern. Eine weitere Ursache für den Registrierungsfehler liegt in der unterschiedlichen Auflösung der Aufnahmen und der damit verbundenen Segmentierungsungenauigkeiten. Wohingegen der Schichtabstand in den CT-Aufnahmen bei ca. 1 mm liegt, liegt dieser mit 3 – 5 mm bei den MRT-Aufnahmen deutlich höher. Dies hat zur Folge, dass die bestimmten Targetmittelpunkte mit einer gewissen Unsicherheit in der Höhe einiger weniger mm versehen sind.

Das rein bildbasierte Verfahren schneidet enttäuschend ab, obwohl die registrierten Aufnahmen zunächst plausibel aussehen und auf den ersten Blick oftmals besser ausgerichtet sind als dies bei dem in dieser Arbeit beschriebenen Verfahren den Anschein hat. Da allerdings prinzipbedingt sämtliches Gewebe bei der Registrierung berücksichtigt wird, werden oft Strukturen aufeinander abgebildet, die zwar ähnlich aussehen, nicht jedoch identisch sind, wie z.B. benachbarte Lymphknoten.

**Qualität der deformierten Aufnahmen:** Neben der Registrierungsgenauigkeit spielt auch die Qualität der deformierten Aufnahmen eine entscheidende Rolle. Insbesondere bei multimodaler Registrierung erzeugen rein bildbasierte Verfahren oft unnatürliche Deformationsfelder, weil für sämtliche Strukturen im Referenzbild korrespondierende Strukturen im Modellbild gefunden werden sollen. In verschiedenen Bildmodalitäten sind aber oftmals unterschiedliche Strukturen zu erkennen.

Um zumindest einen Eindruck von der Qualität der deformierten Aufnahmen zu erhalten, wurde diese für die unterschiedlichen Verfahren bewertet. Dazu begutachtete ein Beobachter (erfahrener Bildverarbeiter) alle deformierten Aufnahmen und beurteilte die Qualität der Aufnahmen. Folgende Kategorien standen zur Auswahl:

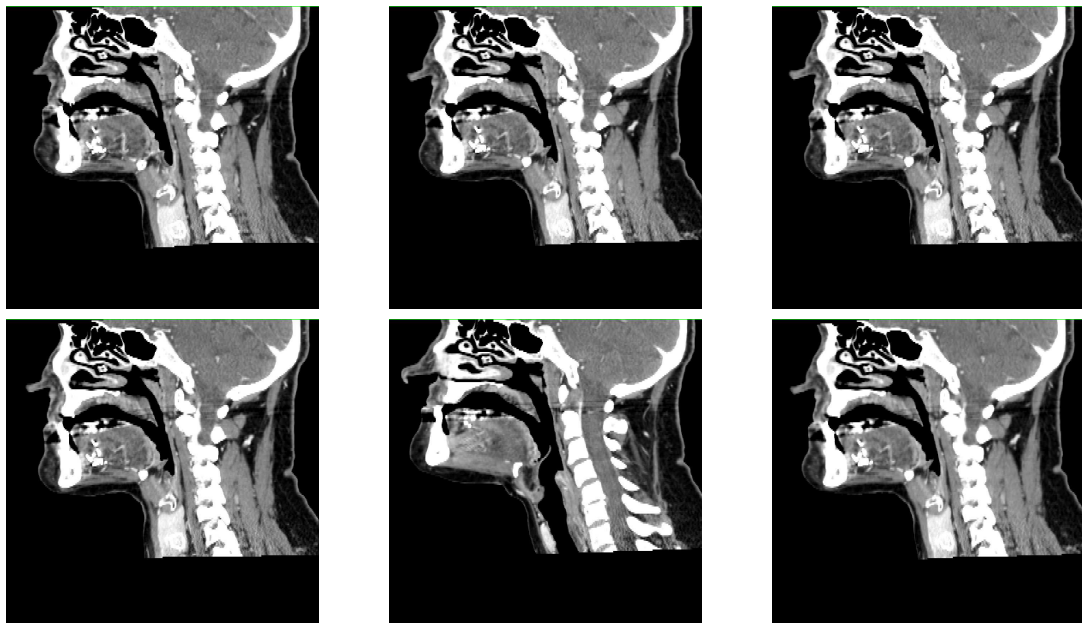
- **Keine erkennbaren Deformationen:** Die Abbildung ist plausibel. Unnatürliche Deformationen sind nicht erkennbar.
- **Leichte Deformationen:** Bei genauerem Hinsehen sind einige wenige unnatürliche Deformationen erkennbar.
- **Starke Deformationen:** Die Aufnahme ist auf den ersten Blick stark unnatürlich deformiert.

Da für diese subjektive Beurteilung keine Referenzdeformation notwendig ist, wurde sie neben den deformierten MRT-Aufnahmen auch für die deformierten CT-Zeitserienaufnahmen durchgeführt. Zum Vergleich wurden wieder die B-Spline deformierten Aufnahmen und die rigide transformierten Aufnahmen herangezogen, wobei bei Letzteren keine Deformationen zu erwarten sind. Um etwaige Vorurteile gegenüber einzelnen Verfahren auszuschließen, wurden die deformierten Aufnahmen unterschiedlicher Verfahren dem Beobachter in zufälliger Reihenfolge präsentiert.

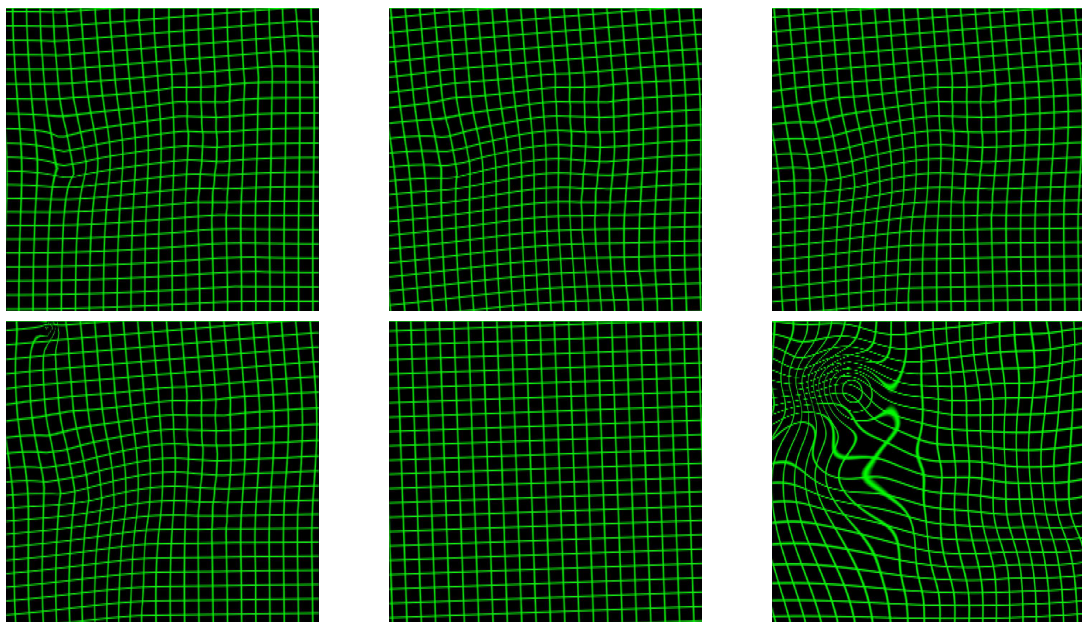
Die Ergebnisse sind in Abbildung 5.20 dargestellt. Einige deformierte Aufnahmen sind exemplarisch in Abbildung 5.21 für den CT-Zeitserienfall und in Abbildung 5.22 dargestellt.

Bei den deformierten CT-Zeitserienaufnahmen zeigt sich, dass die 2-stufige Propagierung der Transformationen zu realistischen Deformationen führt, die von dem Beobachter vergleichbar mit den nicht deformierten Aufnahmen, also den rigide transformierten, bewertet wurden. Die 2-Stufen-Propagierung ist allerdings notwendig. Wird das Transformationsfeld in einem Schritt berechnet, wurden die deformierten Aufnahmen leicht schlechter bewertet. Interessanterweise erzielt die einfache Propagierung der Deformationen eine vergleichbare Qualität wie die 2-stufige Propagierung der Transformationen. Der Grund dafür könnte sein, dass generell die Patienten in den CT-Zeitserienaufnahmen in ähnlichen Wirbelsäulenpositionen aufgenommen wurden, so dass die Deformationen eher gering waren und so die Art der Erzeugung des dichten Deformationsfeldes keine große Rolle spielt. Die Einbeziehung der Körperoberfläche in die Berechnung des Deformationsfeldes führt zu deutlich schlechter wahrgenommenen deformierten Aufnahmen. Der Grund dafür ist, dass das einfache Verfahren zur Oberflächenextraktion insbesondere bei vorhandenen Bildartefakten zu falschen Ergebnissen führt, die dann wiederum unnatürliche Deformationen zur Folge haben. Auch dann, wenn unterschiedliche Bereiche des Kopfes in den beiden Aufnahmen zu sehen sind, entstehen ungewollte Deformationen. Die rigiden Aufnahmen wurden, obwohl Deformationen prinzipbedingt ausgeschlossen sind, nicht komplett ohne Deformation bewertet. Eine Erklärung für dieses überraschende Ergebnis könnte sein, dass vorhandene Bildartefakte vom Beobachter versehentlich als eine unnatürliche Deformation interpretiert wurden. Die mit Abstand schlechtesten Bewertungen werden mit dem B-Spline-Verfahren erzielt. Der Grund dafür ist, dass der gesamte Bildinhalt in den Registrierungsprozess einbezogen wird und nur eine lokale Regularisierung zur Anwendung kommt. Bei Zeitserien kommt hinzu, dass Änderungen in der Anatomie, etwa durch einen chirurgischen Eingriff, dazu führen, dass keine korrekte Lösung des Registrierungsproblems existiert, weil in beiden Aufnahmen nicht dieselben Strukturen zu sehen sind. Die B-Spline-Transformation versucht dennoch für alle Strukturen Korrespondenzen zu finden, was zwangsläufig zu unnatürlichen Deformationen führt. Das in dieser Arbeit vorgestellte Verfahren hat dieses Problem nicht, solange die Anatomieänderungen nicht die Knochen betreffen, weil für das gesamte Weichteilgewebe keine Korrespondenzen gesucht werden.





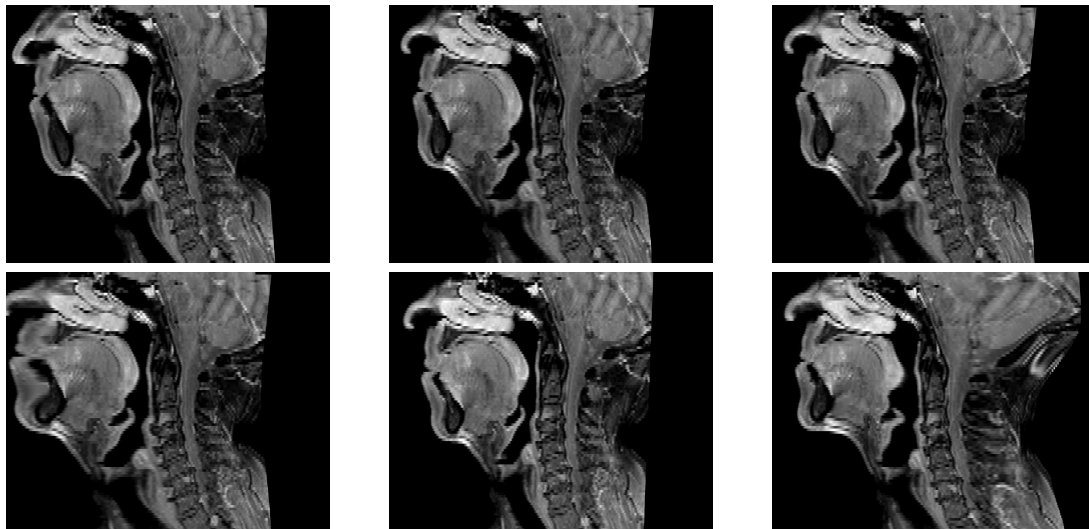
(a) Deformierte Aufnahmen



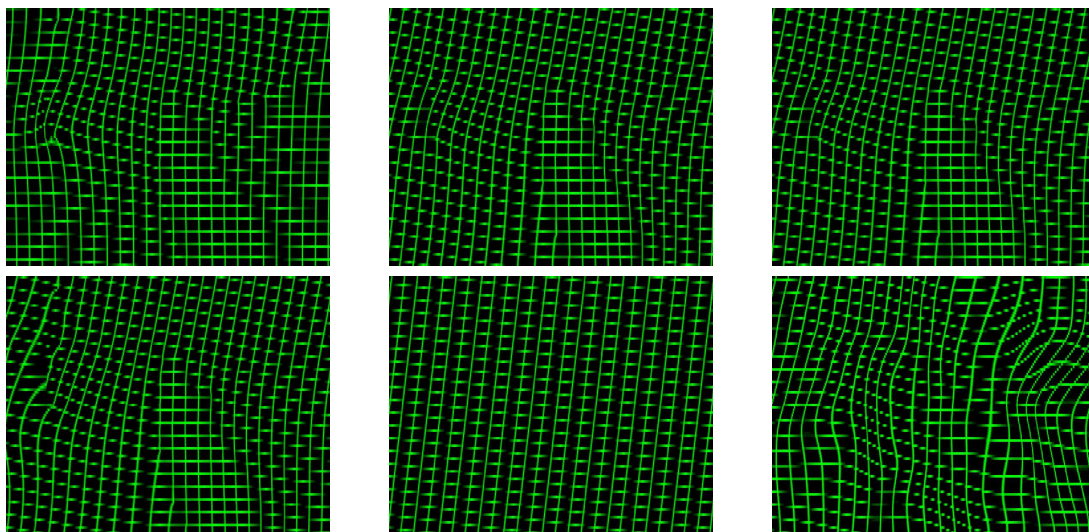
(b) Deformationsfelder

Abbildung 5.21.: Visueller Vergleich der Deformationsfelder von CT-Zeitraufnahmen für verschiedene Arten der Berechnung des Deformationsfeldes an einem Beispiel (obere Reihe: Deformation, Transformation, 2-Stufen; untere Reihe: Oberfläche, Rigide, B-Spline)





(a) Deformierte Aufnahmen



(b) Deformationsfelder

Abbildung 5.22.: Visueller Vergleich der Deformationsfelder von MRT-Aufnahmen für verschiedene Arten der Berechnung des Deformationsfeldes an einem Beispiel (obere Reihe: Deformation, Transformation, 2-Stufen; untere Reihe: Oberfläche, Rigide, B-Spline)

Im multimodalen Fall schneiden die Verfahren in puncto wahrgenommener Qualität der deformierten Aufnahmen ähnlich ab. Eine Ausnahme ist die Propagierung der Deformationen, welche in diesem Fall, wie bei der Entwicklung erwartet, schlechter als die 2-stufige Propagierung der Transformationen abschneidet. Das Einbeziehen der Rotationen in den Deformationsfeldgenerierungsprozess führt also neben einer Verbesserung der Registrierungsgenauigkeit auch zu realistischeren Deformationen. Ein weiterer Unterschied zum CT-Zeitserienfall tritt bei allen Verfahren auf, die das Weichteilgewebe einbeziehen. Darunter fällt das Verfahren, welches die Körperoberfläche einbezieht und das B-Spline-Verfahren. Diese beiden Verfahren schneiden in Bezug auf Qualität des Deformationsfeldes deutlich schlechter ab. Zum einen ist in den MRT-Aufnahmen eine größere Anzahl von Bildartefakten vorhanden. Zum anderen sind die abgebildeten Strukturen in beiden Modalitäten zwar dieselben, werden aber auf sehr unterschiedliche Art und Weise dargestellt, was einen entscheidenden Nachteil für alle Registrierungsverfahren bedeutet, die auf Bildähnlichkeit basieren.

### 5.2.6. Lokal rigide Registrierung von Lymphknoten

Um die Registrierungsgenauigkeit innerhalb der Zielstrukturen, also der Lymphknoten, zu erhöhen wurde die lokal rigide Registrierung, welche in Abschnitt 4.2.8 vorgestellt wurde, auf alle Lymphknoten, die schon in der vorangegangenen Evaluierung der Registrierungsgenauigkeit dienten, als Zielstruktur angewendet. Initialisiert wurde die Transformation in jedem einzelnen Lymphknoten durch Auswertung des dichten Transformationsfeldes, welches durch das Verfahren unter Einbeziehung der Körperoberfläche berechnet wurde, im Schwerpunkt des Lymphknotens. Anschließend wurde der Target-Registration-Error berechnet. Zusätzlich wurde gezählt, in wie vielen Fällen die Registrierungsgenauigkeit unter  $3\text{ mm}$  beträgt.  $3\text{ mm}$  entspricht in etwa dem Schichtabstand der MRT-Aufnahmen und damit auch der maximal erreichbaren Messgenauigkeit des Registrierungsfehlers.

Der durchschnittliche TRE von  $5.14\text{ mm}$  ist mit dem des Kopf-Hals-Registrierungsverfahrens unter Einbeziehung der Körperoberfläche ( $5.05\text{ mm}$ ) vergleichbar. Die Standardabweichung des TREs von  $4.53\text{ mm}$  ist jedoch deutlich höher ( $3.55\text{ mm}$ ). Der Grund dafür ist, dass das lokal rigide Verfahren einen relativ geringen Konvergenzbereich aufweist. Bereits recht gute Registrierungsergebnisse, die durch die Kopf-Hals-Registrierung erzielt wurden, können durch die lokal rigide Registrierung noch weiter verbessert werden, bei schlechten Registrierungsergebnissen tritt eine Verbesserung jedoch nur zufällig ein. Dies schlägt sich auch in der Erfolgsrate nieder, welche sich von nur 27% für das Kopf-Hals-Registrierungsverfahren auf 42% für das lokal rigide Verfahren erhöht.

Aus diesen Ergebnissen kann geschlossen werden, dass das Verfahren bedingt geeignet ist, um die Registrierungsgenauigkeit zu erhöhen. Durch eine nachträgliche visuelle Begutachtung des Registrierungsergebnisses durch einen Experten könnten Verschlechterungen erkannt und herausgefiltert werden. Dies erfordert allerdings Interaktion.

### 5.2.7. Laufzeit

Annähernd der gesamte Registrierungsprozess wurde im Hinblick auf die Laufzeit untersucht. Nur der letzte Schritt, die lokal rigide Registrierung, wurde ausgenommen, weil diese zum einen von der Anzahl der zu untersuchenden Strukturen abhängt und zum anderen vernachlässigbar gegenüber der Gesamtlaufzeit ist. Durchgeführt wurden die Laufzeitmessungen auf einem Intel Core i7-2600K Prozessor mit 3.40 GHz und 8 Prozessorkernen. Auf Parallelisierung wurde teilweise zurückgegriffen.

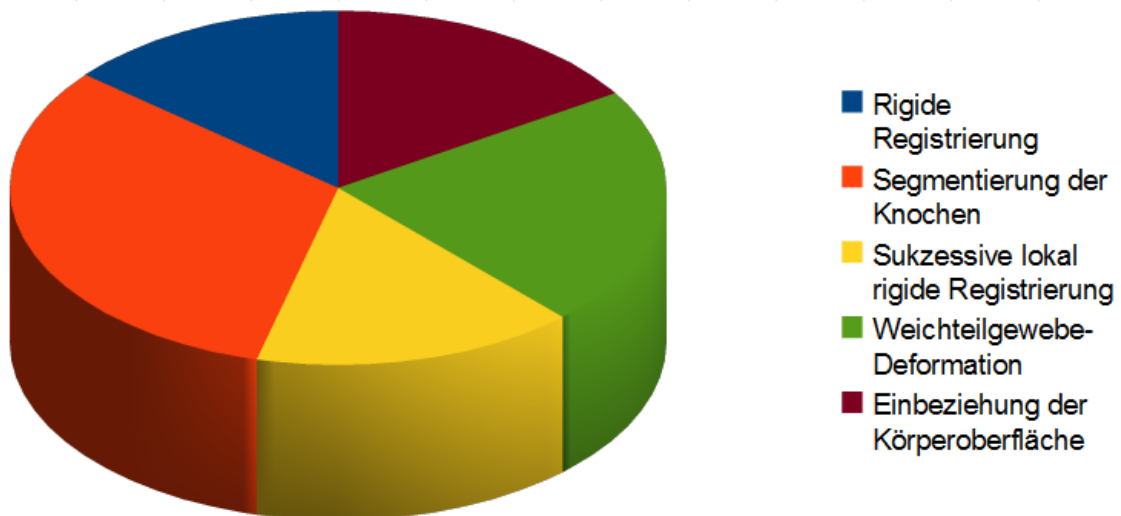


Abbildung 5.23.: Anteil der einzelnen Komponenten an der durchschnittlichen Gesamtlaufzeit von 107 Sekunden für die Registrierung zweier Kopf-Hals-Datensätze

Im Durchschnitt benötigte der gesamte Prozess 107 Sekunden. Wie in Abbildung 5.23 zu sehen ist, sind alle Komponenten zu ähnlichen Teilen an der Gesamtlaufzeit beteiligt. Mit einem Anteil von 32% ist die Segmentierung der Knochen durch den artikulierten Atlas die aufwändigste Komponente.

Im Vergleich dazu benötigte das hierarchische B-Spline-Registrierungsverfahren im Durchschnitt 325 Sekunden zur Registrierung. Das vorgestellte Verfahren ist also trotz der Komplexität der vielen unterschiedlichen Komponenten, welche zum großen Teil nicht parallel ablaufen können, deutlich schneller als das Standardverfahren der deformierbaren Registrierung.

### 5.2.8. Zusammenfassung

Neben der Registrierungsgenauigkeit und der Qualität der erzeugten Deformationsfelder des in dieser Arbeit vorgestellten Registrierungsverfahrens wurde auch jede einzelne Teilkomponente evaluiert. Dabei wurde sowohl der multimodale als auch der Zeitserien-Fall abgedeckt. Die Testaufnahmen wurden im klinischen Alltag aufgenommen, der Aufnahmebereich variierte und es sind zum Teil starke Bildartefakte vorhanden. Der Deformationsgrad der Aufnahmen deckt die typischerweise vorkommende Bandbreite ab. Die Registrierungsgenauigkeit wurde durch den Target Registration Error (TRE) gemessen, wobei Zentren von Lymphknoten als Zielstrukturen dienten. Im Folgenden werden die wichtigsten Ergebnisse zusammengefasst:

- **Registrierungsgenauigkeit:** Der durchschnittliche TRE beträgt 5.05 mm. Gegenüber rigider Registrierung entspricht dies einer Verbesserung von 39%. Gegenüber dem Standardverfahren der nicht-rigiden bildbasierten Registrierung, der B-Spline Registrierung, ergibt sich eine Verbesserung von 37%. Dieser Gewinn an Registrierungsgenauigkeit spricht für die Einbeziehung von globalem Wissen über die Anatomie in den Registrierungsprozess. Die Registrierungsgenauigkeit steigt mit der Nähe an den Knochen und der Ähnlichkeit der Wirbelsäulenposition in den Aufnahmen.

- **Deformationsfeldqualität:** Die Qualität der erzeugten Deformationsfelder des in dieser Arbeit vorgestellten Verfahrens wurde deutlich besser bewertet als die durch ein bildbasiertes Registrierungsverfahren erzeugten Deformationsfelder. Insbesondere ist das Verfahren äußerst robust gegenüber Bildartefakten und Änderungen in der Anatomie.
- **Rigide Registrierung des Kopfes:** In 89% aller Fälle erzielt die vollautomatische rigide Registrierung das gewünschte Registrierungsergebnis, in weiteren 2% gibt es geringfügige Ungenauigkeiten. Die größten Probleme ergeben sich bei stark abgeschnittenen Aufnahmen. Im Vergleich dazu erzielt einfache Mutual Information basierte rigide Registrierung in nur 10% aller Fälle das gewünschte Ergebnis.
- **Erfolgsrate der Knochensegmentierung:** Die Anpassung des artikulierten Atlas an einen ungesesehenen Datensatz verläuft durch das neuartige adaptive Gradientenabstiegsverfahren in 86% aller Fälle erfolgreich. Hauptursache für einen Fehlschlag ist das Verrutschen des Atlas um einen Wirbel. Bei Verwendung des einfachen Gradientenabstiegs tritt dieser Fall deutlich häufiger auf. Die Anpassung verläuft damit in nur 70% aller Fälle erfolgreich.
- **Segmentierungsgenauigkeit der Knochen:** Bei der Segmentierung der Knochen wird ein durchschnittlicher DSC von 0.75 und eine ASD von 1.60 mm erreicht.
- **Registrierung der Knochen:** Die unabhängige lokal rigide Registrierung ist zur Registrierung der Knochen ungeeignet (Erfolgsrate < 25%). Die sukzessive lokale Registrierung der Knochen schneidet insbesondere im multimodalen Fall deutlich besser ab und erzielt eine Erfolgsrate von 90%. Mit einer Erfolgsrate von 94% im CT-Zeitserienfall und 95% im multimodalen Fall werden die besten Ergebnisse allerdings vom (personalisierten) artikulierten Atlas erzielt. Globale Regularisierung ist also tatsächlich von Vorteil.
- **Berechnung des Deformationsfeldes:** Die zusätzliche Propagierung des Rotationsanteils der Transformationen der Knochen führt gegenüber der einfachen Propagierung der Deformationen zu einer zu einer höheren Registrierungsgenauigkeit (+2%) und zum anderen zu realistischen Deformationen. Dieser Effekt tritt allerdings nur ein, wenn das Transformationsfeld in einem 2-stufigen Prozess berechnet wird, wobei die erste Stufe sich auf den Bereich innerhalb des Körpers beschränkt.
- **Einbeziehung der Körperoberfläche:** Die Einbeziehung der Körperoberfläche in den Registrierungsprozess führt zu einer Erhöhung der durchschnittlichen Registrierungsgenauigkeit von ca. 6%. Im Gegenzug ist das Verfahren deutlich anfälliger für Bildartefakte. Die Folge sind unnatürliche Deformationen. Dieser Effekt kommt insbesondere bei der multimodalen Registrierung zum Tragen.
- **Lokal rigide Registrierung der Lymphknoten:** Die Erfolgsrate der Registrierung der Lymphknoten erhöht sich durch das lokal rigide Lymphknotenregistrierungsverfahren von 27% auf 42%. Der Konvergenzbereich ist jedoch ziemlich gering, wodurch sich die Registrierungsgenauigkeit im Durchschnitt sogar leicht verschlechtert (−2%). Für die nicht erfolgreichen Registrierungen ist ein manueller Eingriff nötig.
- **Laufzeit:** Das vorgestellte Verfahren benötigt in der Regel 1 – 2 Minuten zur kompletten Registrierung des Kopf-Hals-Bereichs. Dies ist weniger als die Hälfte der benötigten Laufzeit eines bildbasierten Standardverfahrens.

Das in dieser Arbeit vorgestellte Registrierungsverfahren liefert in puncto Registrierungsgenauigkeit, Qualität der Deformationsfelder und Laufzeit deutlich bessere Ergebnisse als bildbasierte Registrierungsverfahren. Insbesondere ist die Anfälligkeit gegenüber Bildartefakten deutlich geringer.

Verbesserungsbedarf besteht vor allem bei der Segmentierung der Knochen. Eine Vergrößerung der Anzahl an Trainingsdaten könnte dazu behilflich sein.



## 6. Anwendungen

Die in dieser Arbeit vorgestellten Methoden zur Radialstrahl-basierten Segmentierung und zur multirigidn Registrierung wurden zur Merkmalsextraktion aus multimodalen Zeitreihen zur computer-gestützten Vorhersage des Krankheitsverlaufs von Patienten mit Kopf-Hals-Karzinom entwickelt. Aufgrund der Generalisierbarkeit der verfolgten Ansätze lassen sich die im Rahmen dieser Arbeit ent-standenen Technologien oder Teile davon in leichter Abwandlung auch in anderen unterschiedlichen Anwendungsszenarien der medizinischen Bildverarbeitung verwenden. Für einige Anwendungsszenari-en ist dies bereits geschehen, andere Bereiche sind denkbar. Dieses Kapitel soll einen kurzen Überblick über die Anwendungen der entwickelten Methoden geben.

### 6.1. Lymphknotenanalyse einer multimodalen Zeitreihe zur Krankheitsverlaufsvorhersage

Die manuelle Analyse einzelner Lymphknoten in einer einzigen Modalität zu einem einzigen Zeit-punkt hat folgende Nachteile (vgl. Abschnitt 1.1): ungenaue Größenbestimmung, hohe Inter- und Intra-Observer-Variabilität, wenig quantitative Aussagen, geringes Ausnutzen der komplementären Information aus unterschiedlichen Bildmodalitäten und nur rudimentäre Verlaufsanalyse.

Auf Grundlage der in dieser Arbeit entwickelten Methoden zur Registrierung und Segmentierung von Lymphknoten wurde ein interaktives Softwaretool erstellt, welches es ermöglicht, aus zu unterschiedli-chen Zeitpunkten akquirierten CT- und MRT-Aufnahmen, klinisch relevante Merkmale zu extrahieren und an einen Server zu schicken, der diese mit klinischen und genomischen Merkmalen integriert und daraus den Krankheitsverlauf, oder genauer die Wahrscheinlichkeit eines Rezidivs, berechnet [3][4]. Das Schaubild in Abbildung 6.1 veranschaulicht dies. Jeder klinisch relevante Lymphknoten wird da-bei durch ein multimodales, multitemporales parametrisches Modell repräsentiert. Im Gegensatz dazu kommen bei dem Primärtumor keine Informationen über mehrere Zeitpunkte hinweg zum Einsatz, weil dieser nach Diagnose in der Regel chirurgisch entfernt wird. In der klinischen Praxis werden auch Lymphknoten sicherheitshalber großflächig entfernt, was aber erhebliche Nebenwirkungen und Ein-bußen in der Lebensqualität nach sich zieht. Durch die Bestimmung des individuellen Risikos durch die computergestützte Vorhersage des Krankheitsverlaufs kann aber in vielen Fällen das präventive Entfernen der Lymphknoten durch eine Verlaufskontrolle ersetzt werden.

**Eingabedaten:** In dieser konkreten Anwendung werden Aufnahmen zunächst bei Diagnose und dann nach Therapie alle sechs Monate aufgenommen. Folgende Bildmodalitäten kommen in der Regel zum Einsatz:

- CT, nativ
- CT, kontrastverstärkt

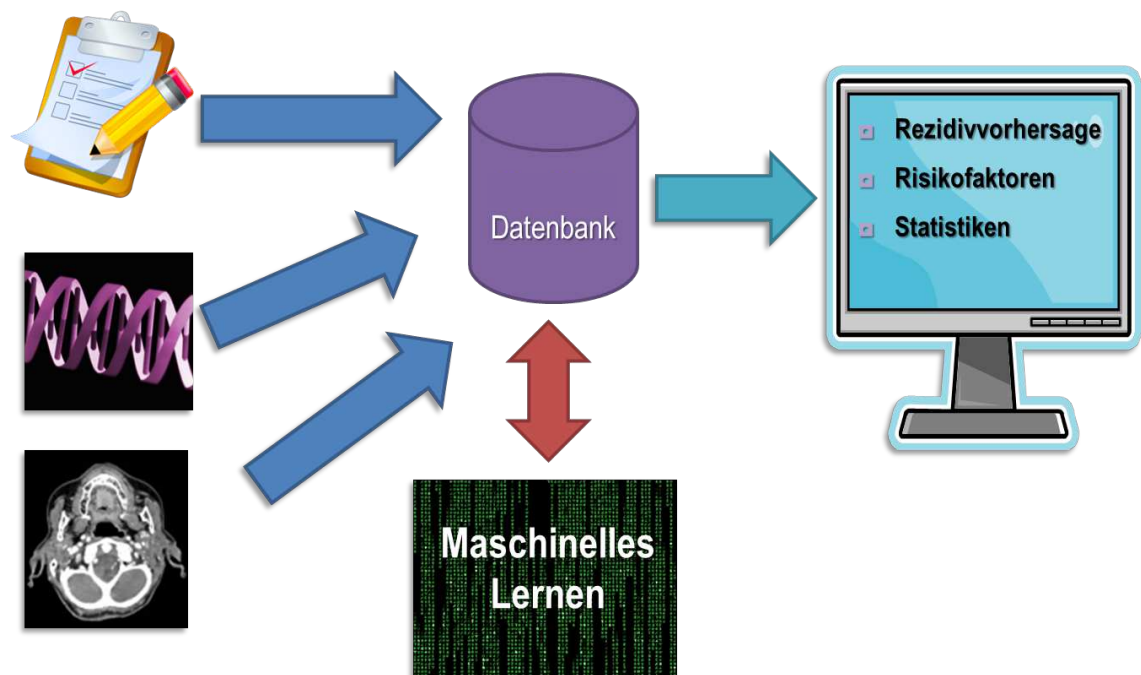


Abbildung 6.1.: Integration der bildbasierten mit klinischen und genomischen Merkmalen zur computergestützten Vorhersage des Krankheitsverlaufs von Patienten mit Kopf-Hals-Karzinom

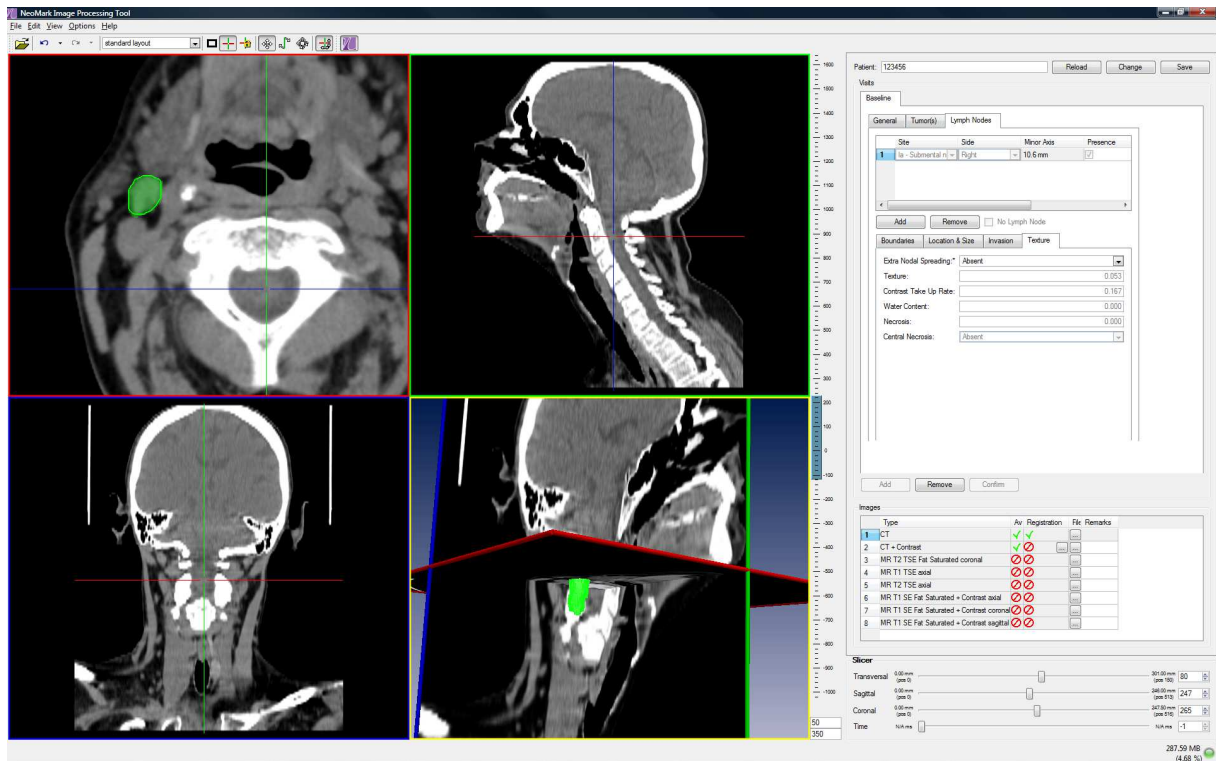
- MRT, koronal, T2 gewichtet, Turbo Spin Echo, Fett-unterdrückt
- MRT, axial, T1 gewichtet, Turbo Spin Echo
- MRT, axial, T2 gewichtet, Turbo Spin Echo
- MRT, axial, T1 gewichtet, Fett-unterdrückt, kontrastverstärkt
- MRT, koronal, T1 gewichtet, Fett-unterdrückt, kontrastverstärkt
- MRT, sagittal, T1 gewichtet, Fett-unterdrückt, kontrastverstärkt

Die beiden CT-Aufnahmen haben einen Schichtabstand von ca. 1 mm und decken den gesamten Kopf-Hals-Bereich ab. Die MRT-Aufnahmen beschränken sich auf den relevanten Bereich des Primärtumors und der Lymphknoten. Sowohl die CT-Aufnahmen, als auch die MRT-Aufnahmen sind für sich jeweils gut registriert, weil alle Aufnahmen am Stück im selben Scanner durchgeführt wurden. Zwischen CT- und MRT-Aufnahmen, als auch bei Aufnahmen zu unterschiedlichen Zeitpunkten ist eine deformierbare Registrierung jedoch Voraussetzung für eine gemeinsame Weiterverarbeitung.

**Bildverarbeitung:** Die in dieser Arbeit entstandenen Innovationen der medizinischen Bildverarbeitung sind die wichtigsten Grundbausteine des Softwaretools zur Merkmalsextraktion (siehe Abbildung 6.2(a)). Diese Bausteine werden durch die in Abbildung 6.2(b) dargestellten Bildverarbeitungskette zusammengefügt.



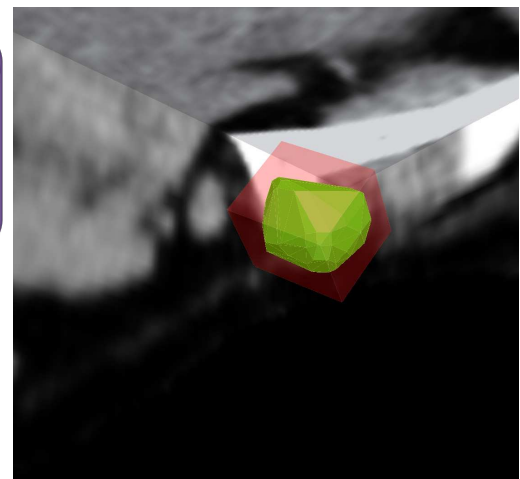
## 6.1. Lymphknotenanalyse einer multimodalen Zeitreihe zur Krankheitsverlaufsvorhersage



(a) Screenshot des Softwaretools



(b) Bildverarbeitungskette basierend auf der in dieser Arbeit entwickelten Methoden zur Bildregistrierung und Segmentierung



(c) Umschließender Quader mit minimalem Volumen zur Berechnung der 3D-Abmessungen

Abbildung 6.2.: Softwaretool zur Extraktion von Merkmalen aus multimodalen Zeitreihen zur computergestützten Vorhersage des Krankheitsverlaufs von Patienten mit Kopf-Hals-Karzinom

| Größte Abmessung | Infiltration   | T-Klassifikation |
|------------------|--|------------------|
| -                | Flügel-Gaumen-Grube, Masticator Space, Karotiden oder Schädelbasis | T4B              |
| -                | andere   | T4A              |
| > 40 mm          | keine  | T3               |
| > 20 mm          | keine  | T2               |
| kleiner          | keine  | T1               |

Tabelle 6.1.: Klassifikation des Tumors aufgrund der 3D-Abmessungen und Infiltrationseigenschaften

Zunächst wird das in Abschnitt 4.2 beschriebene Registrierungsverfahren auf die multimodalen Aufnahmen des initialen Zeitpunkts angewendet. Somit sind alle MRT-Aufnahmen an den CT-Aufnahmen ausgerichtet. Dies ermöglicht dem Radiologen schon eine Betrachtung der fusionierten Bilddaten.

Anschließend wird der Primärtumor in einer beliebigen Modalität durch das interaktive Radialstrahl-basierte Verfahren (siehe Abschnitt 4.1.8) segmentiert und ggf. manuell korrigiert. Anschließend wählt der Anwender die klinisch relevanten Lymphknoten durch je einen Klick in das ungefähre Zentrum des Lymphknotens aus. In der CT-Aufnahme werden diese dann durch das in Abschnitt 4.1.3 vorgestellte Radialstrahl-basierte Lymphknotensegmentierungsverfahren segmentiert. Bevor multimodale Modelle der nun definierten Zielorgane erstellt werden können, wird die Registrierungsgenauigkeit noch durch das in Abschnitt 4.2.8 vorgestellte lokal-rigide Registrierungsverfahren erhöht.

Die CT-Aufnahme des ersten Zeitpunkts dient als Referenz für alle folgenden Zeitpunkte. Werden Aufnahmen während der Verlaufskontrolle hinzugefügt, werden diese zunächst an der Referenzaufnahme wiederum durch das multi-rigide Registrierungsverfahren aus Abschnitt 4.2 ausgerichtet. Auf diese Art und Weise entsteht eine Multikanalaufnahme, die unterschiedliche Bildmodalitäten und Zeitpunkte umfasst. Die Lymphknoten des initialen Zeitpunkts werden nun im folgenden Zeitpunkt wiederum durch das Radialstrahl-basierte Verfahren aus Abschnitt 4.1.3 segmentiert. Darüber hinaus können Lymphknoten, welche zuvor nicht für klinisch relevant befunden wurden, zusätzlich aufgenommen werden. Nun liegt für jeden Lymphknoten eine multimodale Zeitreihe vor, welche später noch um weitere Zeitpunkt bzw. Bildmodalitäten erweitert werden kann.

**Merkmalsextraktion:** Aus der multimodalen Zeitreihe werden dann charakteristische Merkmale aus Geometrie und Textur extrahiert. Diese werden dann noch durch zusätzliche durch den Anwender eingegebene Informationen angereichert. Im Folgenden werden einige dieser Merkmale und deren Art der Berechnung beschrieben:

- **Volumen:** Um das Volumen einer Struktur zu bestimmen, werden die enthaltenen Voxel gezählt und mit dem Volumen eines Voxels multipliziert.
- **3D-Abmessungen:** Die 3D-Abmessungen sind näherungsweise durch die Abmessungen des umschließenden Quaders mit minimalem Volumen gegeben (siehe Abbildung 6.2(c)). Dieser wurde ausgehend von der konvexen Hülle der Struktur durch den in [156] beschriebenen Algorithmus berechnet.
- **Tumorklassifikation:** Aufgrund der durch die 3D-Abmessungen gegebenen Größe und manuell eingegebenen Infiltrationseigenschaften wird der Tumor nach den TNM-Kriterien klassifiziert [23]. Tabelle 6.1 beschreibt die Klassifikation.

- **Stadium:** Das Stadium setzt sich aus der Tumorklassifikation (siehe oben), dem Vorhandensein von Lymphknotenmetastasen und dem Vorhandensein von Fernmetastasen zusammen.
- **Wachstumsrate:** Die Wachstumsrate  $w_{t_1, t_2}$  zwischen unterschiedlichen Zeitpunkten  $t_1$  und  $t_2$  einer Struktur wird aus den entsprechenden Volumina  $V_{t_1}$  und  $V_{t_2}$  berechnet:  $w_{t_1, t_2} = \frac{V_{t_2} - V_{t_1}}{V_{t_1}}$ .
- **Wassergehalt:** Der Wassergehalt kann aus den T1- und T2-gewichteten MRT-Aufnahmen bestimmt werden. Dieses ist in T2-gewichteten MRT-Aufnahmen hell dargestellt, während es in T1-gewichteten Aufnahmen dunkel dargestellt ist. Ein quantitatives Maß dafür ist das Produkt der positiven Abweichung von der mittleren Intensität in der T2-gewichteten Aufnahme und der negativen Abweichung von der mittleren Intensität in der T1-gewichteten Aufnahme. Durch die Exponentialfunktion wird dieses stets positive Produkt auf das Intervall  $[0, 1]$  beschränkt.
- **Kontrastmittelaufnahme:** Die Kontrastmittelaufnahme wird aus der mittleren quadratischen Abweichung der kontrastverstärkten und nativen CT-Aufnahme innerhalb der Struktur berechnet. Diese wird wiederum durch die Exponentialfunktion auf das Intervall  $[0, 1]$  beschränkt.
- **Nekrotisches Gewebe:** Die Menge des nekrotischen Gewebes ist durch den Quotienten der als nekrotisch klassifizierten Voxel und allen Voxel einer Struktur beschrieben. Nekrotisches Gewebe ist durch einen erhöhten Wassergehalt charakterisiert, welcher sich durch die weiter oben beschriebene Vorgehensweise berechnen lässt. Ist der Wassergehalt für einen Voxel oberhalb einer bestimmten Schwelle, wird dieser als nekrotisch klassifiziert.

**Semantische Bildsuche:** Da alle Aufnahmen zusammen mit den extrahierten Merkmalen in einer Datenbank vorliegen, wurde zusätzlich eine semantische Suche nach medizinischen Bilddaten implementiert. So kann der Anwender sich alle Fälle anzeigen lassen, die bspw. regionale Lymphknotenmetastasen aufweisen. Auf diese Art und Weise können unterschiedliche Patientendaten sehr komfortabel miteinander verglichen werden, was z.B. bei der Diagnose helfen kann, aber auch klinische Studien vereinfacht.

**Klinischer Einsatz:** Die in dieser Arbeit vorgestellten Technologien zur Bildregistrierung und Lymphknotensegmentierung wurden implementiert und in ein open-source Softwaretool für die medizinische Bildverarbeitung (MITK [157]) integriert. Es ermöglicht die gerade Beschriebene Anwendung, die semi-automatische Merkmalsextraktion von Lymphknoten und Tumoren aus multimodalen Zeitreihen zur Krankheitsverlaufsvorhersage eines Kopf-Hals-Karzinoms. Dieses Softwaretool ist im Einsatz an dem Universitätsklinikum Parma, Italien und dem MD-Anderson Krebszentrum in Madrid, Spanien.



Abbildung 6.3.: Segmentierung von Lebertumoren durch das Radialstrahl-basierte Verfahren in CT-Aufnahmen

## 6.2. Weitere Anwendungen

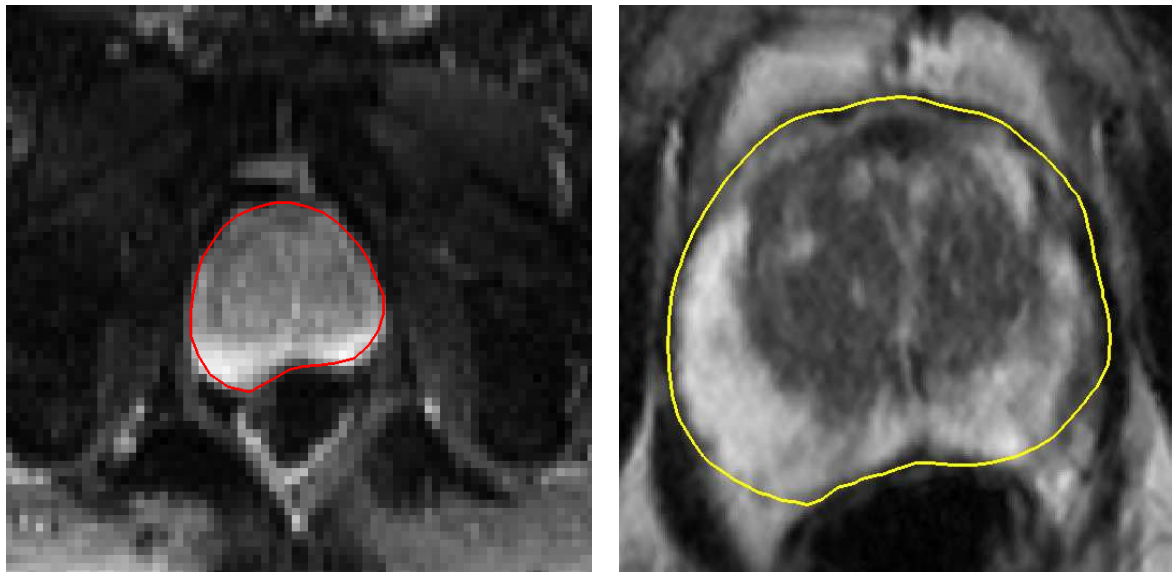
### 6.2.1. Segmentierung von Lebertumoren

Das in Kapitel 4.1.3 vorgestellte Radialstrahl-basierte Segmentierungsverfahren, welches im Rahmen dieser Arbeit für Lymphknoten entwickelt wurde, wurde unter Verwendung der in Abschnitt 4.1.8 vorgestellten interaktiven Erweiterung auch auf Lebertumoren in CT-Aufnahmen angewendet (siehe Abbildung 6.3) und in [12] veröffentlicht. Auch diese Strukturen sind rundlich und haben eine hohe Form- und Größenvarianz. Das umliegende Gewebe, das Leberparenchym, ist jedoch in der Regel etwas homogener. Eine Besonderheit, die bei der Segmentierung beachtet wurde, ist, dass Lebertumore auch am Rand der Leber vorkommen. Dort herrscht oft ein starker Bildgradient, welcher nicht mit dem oft schwächeren Bildgradienten der Tumorgrenze verwechselt werden darf. Als Bildkostenfunktion kommt daher eine Kombination aus richtungsgewichteter logarithmischer Bildgradientenstärke und kumulierten Bildintensitäten zum Einsatz. Zur Evaluation wurde das Verfahren auf 24 Tumore in 3 CT-Datensätzen angewendet. Ein Anwender war bei allen Tumoren in der Lage, in durchschnittlich 15.2 Sekunden zu einem zufriedenstellenden Segmentierungsergebnis zu gelangen.

### 6.2.2. Segmentierung der Prostata aus multiparametrischen MRT-Aufnahmen

Auch bei der Prostata handelt es sich um eine in etwa runde Struktur. Die Segmentierung aus T2-gewichteten MRT-Aufnahmen stellt eine große Herausforderung dar, da das Innere inhomogen ist und sich in der nahen Umgebung der Prostata Bereiche mit ähnlicher Bildcharakteristik befinden. Modellbasierte Verfahren (z.B. [158]) sind durch die Verwendung von globalem Formwissen dennoch in der Lage, gute Ergebnisse zu erzielen. Zum Training des Formenraums und des Aussehens werden allerdings umfangreiche annotierte Trainingsdatensätze benötigt. Sind diese nicht vorhanden, bietet das in dieser Arbeit entwickelte Radialstrahl-basierte Segmentierungsverfahren eine Alternative.

In diffusionsgewichteten MRT-Aufnahmen (DWI) zeichnet sich die Prostata hingegen deutlich stärker



(a) Initiale Segmentierung in niedrigauflösender aber kontrastreicher DWI-gewichteter Aufnahme

(b) Verfeinerung der Kontur in hochauflösender T2-gewichteter Aufnahme

Abbildung 6.4.: Segmentierung der Prostata aus multiparametrischen MRT-Aufnahmen durch das Radialstrahl-basierte Segmentierungsverfahren

vom Hintergrund ab, jedoch ist die räumliche Auflösung gering. Daher wird das Radialstrahl-basierte 3D-Segmentierungsverfahren zunächst auf die Prostata in der DWI-MRT-Aufnahme angewendet. Parallel dazu werden axiale, koronale und sagittale Schichten von T2-gewichteten Aufnahmen durch ein Superresolution-Verfahren [159] zu einer hochauflösenden T2-gewichteten MRT-Volumen fusioniert. An dieses Volumen wird dann die zuvor erhaltene grobe Kontur rigide angepasst. Als Anpassungskriterium dient dabei die mit einem Laplace-Filter gefaltete Version der Superresolution-T2-Aufnahme, wodurch die linienartigen Kanten der Prostatagrenzen hervorgehoben werden. Abschließend wird wiederum das Radialstrahl-basierte Segmentierungsverfahren angewendet, diesmal auf den gefilterten T2-Daten, wobei Orten nahe der rigide transformierten DWI-Kontur geringe bildbasierte Kosten zugeordnet werden. Abbildung 6.4 zeigt ein exemplarisches Segmentierungsergebnis.

### 6.2.3. Lymphoma-Therapieerfolgsbestimmung in PET-CT-Zeitserienaufnahmen

Patienten mit einem malignem Lymphom, also einer bösartigen Veränderung des lymphatischen Gewebes, werden meist durch eine systemische Chemotherapie behandelt. Unterschiedliche Patienten sprechen auf unterschiedliche Stoffe unterschiedlich stark an. Um den Therapieerfolg zu bestimmen, wird neben der Größenveränderung durch die RECIST-Kriterien [26] auch die Glukosestoffwechselveränderung durch PET-Aufnahmen mit FDG-Tracer im Laufe der Zeit gemessen [160]. Aus Zeitgründen geschieht dies in der klinischen Praxis jedoch nur für einige wenige Lymphknoten in einer repräsentativen Schicht. Idealerweise sollte aber das Gesamttumolvolumen betrachtet werden.

Das im Rahmen dieser Arbeit entwickelte Verfahren zur Lymphknotensegmentierung diente als Aus-



Abbildung 6.5.: Betrachtung des Gesamttumorzvolumens in Ganzkörper-PET/CT-Aufnahmen zur Bestimmung des Therapieerfolgs bei Lymphompatienten

gangssituation für die Entwicklung eines Softwareprototypen, welcher solch eine volumetrische Betrachtung in registrierten Ganzkörper-PET/CT-Aufnahmen in kurzer Zeit ermöglicht. Saatpunkte für das Segmentierungsverfahren werden automatisch aus den lokalen Maxima des PET-Signals bestimmt. Der Anwender ist dann in der Lage, Saatpunkte zu korrigieren, falsch detektierte Saatpunkte zu entfernen oder Saatpunkte hinzuzufügen. Der Segmentierungsprozess läuft im Hintergrund ab. Für jeden Lymphknoten werden ausgehend von der Segmentierung das maximale und kumulierte PET-Signal sowie das Volumen und die 3D-Achsen bestimmt. Abschließend wird das Gesamttumorzvolumen und das kumulierte PET-Signal für alle Lymphknoten bestimmt und visualisiert (siehe Abbildung 6.5). Für Zeitserienaufnahmen im Laufe der Therapie kann dieser Prozess wiederholt und die Ergebnisse können verglichen werden. Ein geübter Anwender benötigt für die Bestimmung der Parameter aus einer Aufnahme unter 5 Minuten. Eine klinische Evaluierung ist im Gange.

### 6.2.4. Multi-rigide Registrierung der Leber

Die in dieser Arbeit vorgestellte Art der Deformationsfeldberechnung und die unabhängige lokal rigide Registrierung wurde zur deformierbaren Registrierung von diffusionsgewichteten und anatomischen MRT-Aufnahmen der Leber angewendet. Da hier jedoch keine per-se rigiden Strukturen vorhanden sind, wird einfach die gesamte Aufnahme hierarchisch in verschieden große Teile unterteilt. Jeder Teil wird dann für sich genommen lokal rigide durch die Maximierung der Mutual Information registriert. Da aufgrund von geringem Kontrast einige Registrierungen fehlschlagen, sollen diese aussortiert werden. Dazu wird das lokale Maximum der Mutual Information, in welches das Registrierungsverfahren

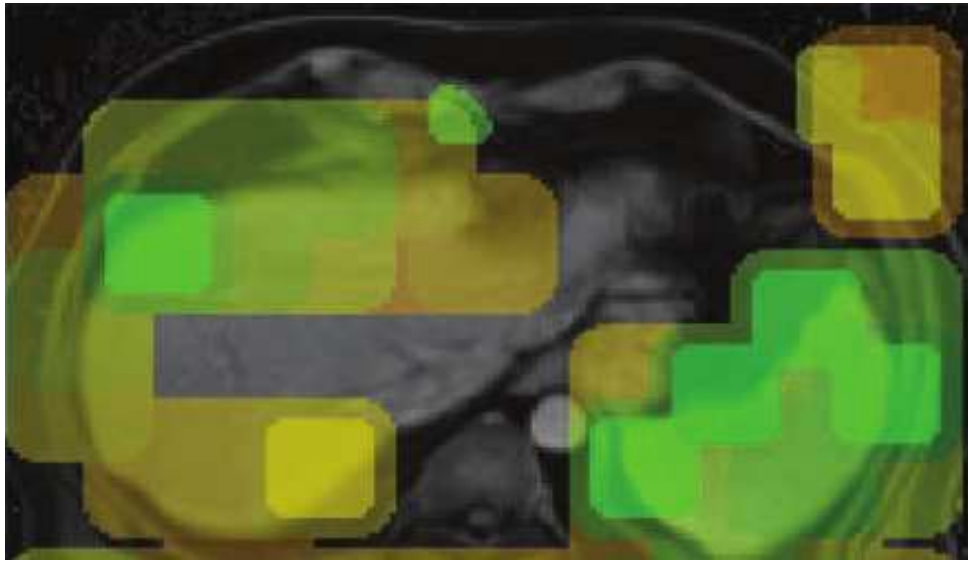


Abbildung 6.6.: Gültige rigide Bildausschnitte zur multi-rigiden Registrierung von diffusionsgewichteten und anatomischen MRT-Aufnahmen der Leber. Die *Spitzheit* des gefundenen lokalen Maximums ist farbkodiert dargestellt, wobei grün der größten Verlässlichkeit entspricht.

konvergiert ist, untersucht. Ein *spitzer Berg* spricht für eine verlässliche Registrierung, wobei die *Spitzheit* durch die Determinante der Hessematrix gegeben ist. So entsteht eine Serie von gültigen Bildausschnitten, für die je eine rigide Transformation gegeben ist (siehe Abbildung 6.6). Ein dichtes Deformationsfeld wird daraus durch das in Abschnitt 4.2.6 vorgestellte Verfahren bestimmt. Weitere Details dazu sind in [7] zu finden.

### 6.2.5. Segmentierung von Knochen in MRT-Aufnahmen

Knochen kann aufgrund des geringen Kontrastes nur sehr schwer in MRT-Aufnahmen segmentiert werden. In CT-Aufnahmen ist der Kontrast hingegen sehr hoch und eine präzise Segmentierung ist möglich. Ein Großteil der in dieser Arbeit vorgestellten Verfahren zur Registrierung von Kopf-Hals-Aufnahmen kann auch zur Segmentierung von Knochen in MRT-Aufnahmen verwendet werden, solange auch eine CT-Aufnahme desselben Patienten vorliegt. Zunächst werden die Knochen durch den in Abschnitt 4.2.2 vorgestellten artikulierten Atlas segmentiert. Darauf aufbauend findet dann eine Personalisierung des artikulierten Atlas statt. Dieser, in Abschnitt 4.2.5 beschriebene, personalisierte artikulierte Atlas wird dann an die MRT-Aufnahme angepasst. In einem letzten Schritt werden dann die Segmentierungen von der CT-Aufnahme in die MRT-Aufnahme übertragen. Wurden die Transformationen der Knochen zwischen der CT- und der MRT-Aufnahme korrekt bestimmt, führt diese Art der *Regmentierung* zu einer Segmentierungsgenauigkeit in MRT-Aufnahmen, die exakt der hohen Segmentierungsgenauigkeit in CT-Aufnahmen entspricht, obwohl die Segmentierung von Knochen in MRT-Aufnahmen ein ungleich schwierigeres Problem ist.



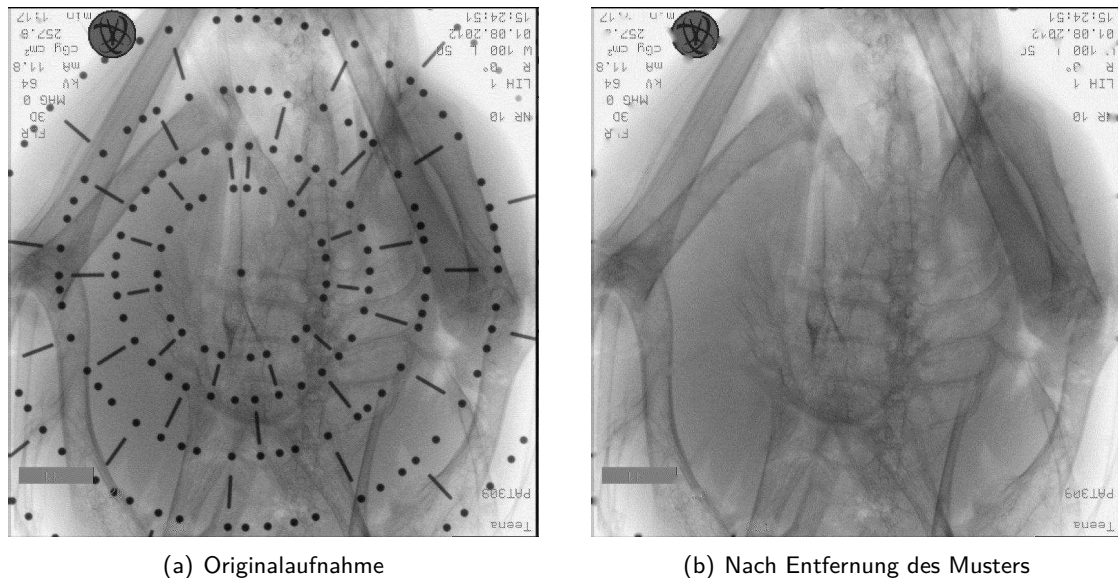


Abbildung 6.7.: Intraoperative Röntgenaufnahmen zur Navigation in der Bronchoskopie mit Hilfe eines Musters (Abbildungen aus [162])

### 6.2.6. Entfernung von Markern aus Fluoroskopieaufnahmen

Zur Erleichterung der Navigation in der Bronchoskopie kann Information aus einer Präoperativen CT-Aufnahme in eine Intraoperative Röntgenaufnahme oder von Röntgenaufnahme in die CT-Aufnahme übertragen werden [161]. Dazu müssen allerdings die Aufnahmeparameter des C-Bogens (Kameraparameter sowie Position und Orientierung) bekannt sein. Eine Variante, die keine Anforderungen an den C-Bogen stellt und somit universell eingesetzt werden kann, basiert auf einem Muster aus Kugeln und Stiften, welches unter dem Patienten auf dem Tisch platziert wird. Durch Projektionsinvarianz von Doppelverhältnissen von Abständen und Winkeln innerhalb des Musters kann die Position und Orientierung eindeutig bestimmt werden. Jedoch kann das in der Röntgenaufnahme sichtbare Muster für einen Anwender störend wirken (siehe Abbildung 6.7(a)). Dieselbe Technologie, welche in dieser Arbeit zur Berechnung des Deformationsfeldes zur Registrierung von Kopf-Hals-Aufnahmen verwendet wurde, wurde auch hier verwendet, um das Muster zu entfernen. Die Segmentierung des Musters bestimmt den Bereich, welcher aus dem Bild *entfernt* und geeignet ersetzt werden soll. Dieser Vorgang wird als *Inpainting* bezeichnet. Als bekannte Information dienen die am Rand der Segmentierung vorkommenden Grauwerte, welche in das Innere der Segmentierung propagiert werden soll. Wie schon bei der Berechnung des Deformationsfeldes, geschieht dies durch die Lösung der Laplacegleichung, wobei am Rand Dirichlet-Randbedingungen zur Anwendung kommen. Die Lösung erzeugt Grauwerte, welche einen maximal glatten Übergang erzeugen (siehe Abbildung 6.7(b)), so dass das Muster keinen störenden Einfluss mehr hat.



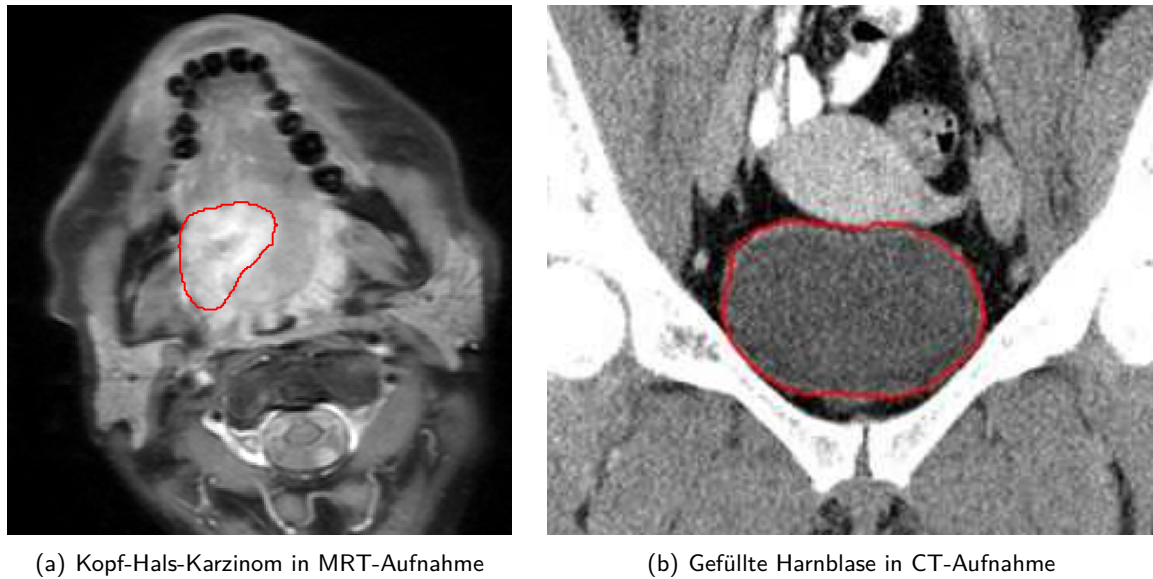


Abbildung 6.8.: Weitere Anwendungsszenarien des Radialstrahl-basierten 3D-Segmentierungsverfahrens

## 6.3. Denkbare Anwendungsgebiete

### 6.3.1. Segmentierung rundlicher Strukturen

Das Radialstrahl-basierte 3D-Segmentierungsverfahren, welches im Rahmen dieser Arbeit zur Segmentierung von Lymphknoten entwickelt wurde, eignet sich auch für eine Reihe anderer in etwa runder Strukturen, welche im menschlichen Körper oft vorkommen. Je nach Bildmodalität und Struktur muss die bildbasierte Kostenfunktion entsprechend angepasst werden. Besonders geeignet ist das Verfahren für sog. Wachstumsstrukturen, also Strukturen, welche von einem Zentrum aus in alle Richtungen mehr oder weniger gleichmäßig wachsen. Das liegt daran, dass das Verfahren sehr gut die hohe Form- und Größenvarianz solcher Strukturen abdecken kann. Beispiele für solche Strukturen sind neben der bereits erwähnten Lymphknoten und Lebertumoren auch Zysten und eine Vielzahl anderer Tumore (siehe Abbildung 6.8(a)).

Neben Wachstumsstrukturen kann das Verfahren auch alternativ zu modellbasierten Verfahren zur Segmentierung von rundlichen Organen eingesetzt werden. Beispiele hierfür sind neben der bereits erwähnten Prostata auch die Harnblase (siehe Abbildung 6.8(b)), das Auge und die Herzkammern.

### 6.3.2. Gemeinsame Segmentierung mehrerer Strukturen durch den artikulierten Atlas

Der in dieser Arbeit vorgestellte artikulierte Atlas wurde zur gemeinsamen Segmentierung des Schädels, des Unterkiefers und der Halswirbel erfolgreich verwendet. Die Besonderheit dabei ist, dass die relativen Lagen der einzelnen Teile aus Trainingsdaten gelernt und in den Segmentierungsprozess eingebunden werden. Dies ist zum einen wie im Kopf-Hals-Bereich bei Knochenstrukturen der Fall,

die über Gelenke miteinander verbunden sind. Beispiele dafür sind die Wirbelsäule, der Brustkorb, die Schulter, die Hand, der Fuß, das Knie, usw. Zum anderen kann dies auch bei der gemeinsamen Segmentierung nah beieinanderliegender Strukturen im Weichteilgewebe von Vorteil sein, insbesondere dann, wenn die zu segmentierende Struktur schlecht zu erkennen ist, und sich hauptsächlich durch die relative Lage zu anderen Strukturen definiert. Beispiele dafür sind:

- **Prostata:** Zur automatischen Extraktion klinischer Parameter für Diagnose, Stadienbestimmung und Verlaufskontrolle von Prostatakrebs wird die Segmentierung der Prostata benötigt. Sie befindet sich zwischen Rektum und Harnblase. Diese Strukturen können den Segmentierungsprozess unterstützen. Möglicherweise bietet auch die Einbeziehung des Beckenknochens einen Mehrwert.
- **Knorpel im Knie:** Der Fortschritt von Arthrose im Kniegelenk kann durch die Messung der Knorpeldicke bestimmt werden. Dies erfordert eine Segmentierung. Eine gemeinsame Segmentierung des Knorpels mit den umliegenden Knochen durch den artikulierten Atlas ist ein vielversprechender Ansatz.
- **Kopf-Hals-Bereich:** Zur Strahlentherapieplanung müssen sowohl Zielorgane als auch Risikoorgane im niedrig aufgelösten nativen CT definiert werden. Im Kopf-Hals-Bereich ist dies eine Vielzahl von Strukturen, wie z.B. das Auge, der Sehnerv, die Speicheldrüsen, der Kehlkopf, der Hirnstamm, die Lymphknotenregionen, usw.. Einige der Strukturen sind kaum erkennbar. Die Lymphknotenregionen sind sogar keine tatsächlichen anatomischen Strukturen, sondern nur durch die Umgebung definiert. Daher muss die Segmentierung auf Grundlage von Wissen über die relative Lage zu gut erkennbaren Strukturen, wie z.B. den Knochen, erfolgen. Der artikulierte Atlas ist in der Lage, so ein Wissen zu repräsentieren.
- **Bauchspeicheldrüse:** Die Bauchspeicheldrüse ist aufgrund des geringen Bildkontrastes ein extrem schwierig zu segmentierendes Organ, welches sich hinter dem Bauchfell, zwischen Magen, Milz, Leber und den großen Blutgefäßen des Bauchraums befindet und teilweise von dem Zwölffingerdarm umfasst wird. In [163] werden zur Segmentierung Relationen zu umliegenden Organen und Gefäßen explizit modelliert. Der artikulierte Atlas könnte dazu als Alternative dienen.

### 6.3.3. Deformierbare Registrierung auf Grundlage rigider Strukturen

Das in dieser Arbeit vorgestellte Registrierungsverfahren lässt sich auch auf andere Körperbereiche jenseits des Kopf-Hals-Bereichs übertragen, ohne dass neue Methodik entwickelt werden muss. Grundvoraussetzung dafür ist, dass die rigiden Bestandteile die Deformation des umliegenden Weichteilgewebes bestimmt. Wie in dieser Arbeit gezeigt, ist die Genauigkeit umso größer, je näher sich das Registrierungsziel an einer der rigiden Stützstrukturen befindet. Scheint ein Körperbereich geeignet zu sein, muss lediglich der artikulierte Atlas neu trainiert werden. Dafür sind Trainingsdatensätze nötig, in denen die rigiden Stützstrukturen annotiert sind. Die Trainingsdatensätze sollten die gesamte Varianz der möglichen Deformation abdecken. Beispiele für Bereiche des menschlichen Körpers für den dies möglich ist, ist der Brustkorb, die Wirbelsäule, die Extremitäten sowie der Hüftbereich. Sogar eine Ganzkörperregistrierung ist bei einer genügend großen Menge an Trainingsdaten denkbar.

## 7. Zusammenfassung & Ausblick

Die Hauptentwicklungen dieser Arbeit sind ein Radialstrahl-basiertes 3D-Segmentierungsverfahren für rundliche Strukturen unterschiedlicher Größe und ein automatisches deformierbares Bildregistrierungsverfahren auf Grundlage rigider Strukturen, welche über den aktuellen Stand der Forschung hinausgehen.

Die Kombination beider Verfahren ermöglicht eine Merkmalsextraktion von Lymphknoten aus einer multimodalen Zeitreihe im Kopf-Hals-Bereich zur computergestützten Vorhersage des Krankheitsverlaufs von Patienten mit Kopf-Hals-Karzinom. Dazu werden zunächst Kopf-Hals-Aufnahmen unterschiedlicher Bildmodalitäten und unterschiedlicher Zeitpunkte automatisch auf Grundlage der Knochen durch das neue Verfahren registriert. Anschließend wird das neue Segmentierungsverfahren auf einen klinisch relevanten Lymphknoten angewendet, um ein multimodales, multitemporales Modell zu erhalten. Vorteile gegenüber konventioneller Betrachtung durch einen Radiologen sind die quantitative Natur der Merkmale und die geringere Intra- und Inter-Observer-Variabilität. Darüber hinaus ermöglicht dies in Kombination mit klinischen und genomischen Merkmalen die Personalisierung der Therapie und der Verlaufskontrolle, wodurch bspw. unnötige chirurgische Eingriffe verhindert werden können.

Da die beiden entwickelten neuartigen Verfahren der medizinischen Bildverarbeitung zwar über die Anwendung, nicht jedoch aus der Sicht der Methodik zusammenhängen, werden diese im Folgenden getrennt in Bezug auf Schwierigkeiten, Stand der Forschung, Prinzip, Evaluation, Beiträgen zur Informatik und Anwendungen zusammengefasst. Die konzeptionellen Unterschiede des Radialstrahl-basierten Segmentierungsverfahrens und des artikulierten Atlas zur Segmentierung der Knochen als integraler Bestandteil des multi-rigiden Registrierungsverfahrens werden anschließend betrachtet. Abschließend wird ein Ausblick auf sinnvolle Anknüpfungspunkte an diese Arbeit gegeben.

### 7.1. Radialstrahl-basierte Segmentierung rundlicher Strukturen

Radialstrahl-basierte 3D-Segmentierung rundlicher Strukturen wurde in dieser Arbeit intensiv anhand der Segmentierung von Lymphknoten in CT-Daten behandelt.

**Schwierigkeiten:** Lymphknoten weisen eine hohe Varianz an Größe, Form und umliegendes Gewebe auf. Darüber hinaus ist der Kontrast zum umliegenden Gewebe oft gering, hohe Bildgradientenstärken in der näheren Umgebung von Lymphknoten können ein Segmentierungsverfahren in die Irre führen und Pathologien – insbesondere nekrotisches Gewebe – verändern das typische Aussehen von Lymphknoten.

| Verfahren           | Jahr | CT | Echtes 3D Verfahren | Segmentierung einzelner Lymphknoten | Ein-Klick-Initialisierung möglich | Verwendung von Wissen über das Aussehen | Verwendung von Formwissen | Globale Segmentierungsstrategie | Berücksichtigung von pathologischen Fällen | Anwendung auf klinische Daten | Vergleich mit Expertensegmentierung |
|---------------------|------|----|---------------------|-------------------------------------|-----------------------------------|---|---------------------------|---------------------------------|--|-------------------------------|-------------------------------------|
| Rogowska (AS)[74]   | 1996 | ✓  | ✗                   | ✓                                   | ✓                                 | ✓                                       | ✗                         | ✓                               | ✗  | ✓                             | ✓                                   |
| Rogowska (LS)[74]   | 1996 | ✓  | ✗                   | ✓                                   | ✗                                 | ✓                                       | ✗                         | ✗                               | ✗  | ✓                             | ✓                                   |
| Rogowska (WT)[74]   | 1996 | ✓  | ✗                   | ✓                                   | ✗                                 | ✓                                       | ✗                         | ✗                               | ✗  | ✓                             | ✓                                   |
| Rogowska (SN)[74]   | 1996 | ✓  | ✗                   | ✓                                   | ✗                                 | ✓                                       | ✓                         | ✗                               | ✗  | ✓                             | ✓                                   |
| Honea[87]           | 1997 | ✓  | ✗                   | ✓                                   | ✗                                 | ✗                                       | ✓                         | ✗                               | ✗  | ✓                             | ✗                                   |
| Honea[88]           | 1999 | ✓  | ✓                   | ✓                                   | ✓                                 | ✗                                       | ✓                         | ✗                               | ✗  | ✗                             | ✗                                   |
| Yan[89]             | 2004 | ✓  | ✗                   | ✓                                   | ✓                                 | ✓                                       | ✓                         | ✗                               | ✗  | ✓                             | ✗                                   |
| Dornheim[81]        | 2006 | ✓  | ✓                   | ✓                                   | ✓                                 | ✓                                       | ✓                         | ✗                               | ✓  | ✓                             | ✓                                   |
| Unal[76]            | 2006 | ✗  | ✗                   | ✓                                   | ✗                                 | ✓                                       | ✓                         | ✗                               | ✗  | ✓                             | ✓                                   |
| Kitasaka[64]        | 2007 | ✓  | ✓                   | ✓                                   | ✓                                 | ✓                                       | (✓)                       | ✗                               | ✗  | ✓                             | ✗                                   |
| Commowick[68]       | 2008 | ✓  | ✓                   | ✗                                   | -                                 | -                                       | -                         | -                               | -  | ✓                             | ✓                                   |
| Maleike[79]         | 2008 | ✓  | ✓                   | ✓                                   | ✗                                 | ✓                                       | ✓                         | ✗                               | ✗  | ✓                             | ✓                                   |
| Gorthi[70]          | 2009 | ✓  | ✓                   | ✗                                   | -                                 | -                                       | -                         | -                               | -  | ✓                             | ✓                                   |
| Moltz[78]           | 2009 | ✓  | ✓                   | ✓                                   | ✗                                 | ✓                                       | ✓                         | ✓                               | ✗  | ✓                             | ✗                                   |
| Zhang[77]           | 2009 | ✗  | ✗                   | ✓                                   | ✗                                 | ✓                                       | ✓                         | ✓                               | ✗  | ✓                             | ✓                                   |
| Dornheim[83]        | 2010 | ✓  | ✓                   | ✓                                   | ✓                                 | ✓                                       | ✓                         | ✗                               | ✓  | ✓                             | ✓                                   |
| Teng[71]            | 2010 | ✓  | ✓                   | ✗                                   | -                                 | -                                       | -                         | -                               | -  | ✓                             | ✓                                   |
| Wang[93]            | 2010 | ✓  | ✓                   | ✓                                   | ✓                                 | ✓                                       | ✓                         | ✓                               | ✗  | ✓                             | ✓                                   |
| Debats (CCRG)[75]   | 2011 | ✗  | ✓                   | ✓                                   | ✓                                 | ✓                                       | ✗                         | ✗                               | ✗  | ✓                             | ✓                                   |
| Debats (ECC)[75]    | 2011 | ✗  | ✓                   | ✓                                   | ✓                                 | ✓                                       | ✗                         | ✗                               | ✗  | ✓                             | ✓                                   |
| Debats (GCS)[75]    | 2011 | ✗  | ✓                   | ✓                                   | ✓                                 | ✓                                       | ✓                         | ✓                               | ✗  | ✓                             | ✓                                   |
| Debats (PSAM)[75]   | 2011 | ✗  | ✓                   | ✓                                   | ✓                                 | ✓                                       | ✓                         | ✗                               | ✗  | ✓                             | ✓                                   |
| Lu[85]              | 2011 | ✓  | ✗                   | ✓                                   | ✗                                 | ✗                                       | ✓                         | ✗                               | ✗  | ✓                             | ✓                                   |
| Barbu[67]           | 2012 | ✓  | ✓                   | ✓                                   | ✓                                 | ✓                                       | ✓                         | ✓                               | ✗  | ✓                             | ✓                                   |
| Beichel[95]         | 2012 | ✓  | ✓                   | ✓                                   | ✓                                 | ✓                                       | ✓                         | ✓                               | ✗  | ✓                             | ✓                                   |
| Feulner[90]         | 2012 | ✓  | ✓                   | ✓                                   | ✓                                 | ✓                                       | ✓                         | ✓                               | ✗  | ✓                             | ✗                                   |
| <b>Diese Arbeit</b> | 2012 | ✓  | ✓                   | ✓                                   | ✓                                 | ✓                                       | ✓                         | ✓                               | ✓  | ✓                             | ✓                                   |

Tabelle 7.1.: Zusammenfassung von Eigenschaften des in dieser Arbeit entwickelten Lymphknoten-segmentierungsverfahrens mit dem Stand der Forschung (vgl. Tabelle 3.1).

**Stand der Forschung:** Seit 1996 werden zur Segmentierung von Lymphknoten zahlreiche Verfahren unterschiedlicher Kategorien in der Literatur vorgestellt. Jedes Verfahren verwendet eine unterschiedliche bildbasierte Kostenfunktion, deren Parameter jeweils nur informell bestimmt wurden. Ein Vergleich verschiedener Kostenfunktionen ist nicht vorhanden. Darüber hinaus ist keines der existierenden Verfahren in der Lage, alle beschriebenen Schwierigkeiten zu adressieren. Ein Vergleich des in dieser Arbeit entwickelten Verfahrens mit existierenden Verfahren ist in Tabelle 7.1 zu finden.

**Prinzip:** Ausgehend von einem manuell vorgegebenen Saatpunkt werden Strahlen radial in alle möglichen Richtungen gleichverteilt geschickt und ein Optimierungsverfahren bestimmt für jeden Strahl unter Einbeziehung von Bildinformation und lokalem Formwissen den bestmöglichen Radius und somit eine Segmentierung. Zum ersten Mal wurden unterschiedliche Kostenfunktionen und Optimierungsverfahren verglichen und optimale Parameter bestimmt. Das globale Optimierungsverfahren Graph-Cut erzielt basierend auf einer Kombination aus der richtungsgewichteten logarithmischen Bildgradientenstärke und kumulierten Intensitäten die besten Ergebnisse. Darauf aufbauend wurden Erweiterungen vorgestellt, welche die Abhängigkeit vom Saatpunkt verringert, die Menge an unterstützten Formen vergrößert und eine Erhöhung der Genauigkeit durch Interaktion ermöglicht.

**Evaluation:** Die Evaluierung erfolgte anhand von 49 repräsentativen Lymphknoten. Trotz der prinzipbedingten Beschränkung auf sternkonvexe Formen sind Radialstrahl-basierte Segmentierungsverfahren in der Lage, die in der Praxis auftretende Formvarianz abzudecken. Mit einer durchschnittlichen Oberflächendistanz von nur 0.46 mm ist die Segmentierungsgenauigkeit im Bereich der manuellen Expertensegmentierung (+2%) und deutlich besser als andere Radialstrahl-basierte (−41% bzw. −25%) und modellbasierte (−29%) Verfahren. Im Vergleich zum Stand der Forschung werden insbesondere stark verrauschte Aufnahmen, pathologische Lymphknoten und Lymphknoten mit geringem Bildkontrast deutlich besser unterstützt, wobei bei letzteren nach wie vor die größten Herausforderungen liegen. Die Inter-Observer-Variabilität zur Volumenbestimmung ist mit 5.9% deutlich geringer als bei der manuellen Bestimmung des Volumens (15.9%). Dadurch konnte nachgewiesen werden, dass dieser erwartete Vorteil computergestützter Verfahren tatsächlich eintritt. Auf aktueller Hardware benötigt die Segmentierung einige wenige Sekunden.

**Eigene Beiträge:** Der Fokus dieser Arbeit lag bei der Entwicklung der Lymphknoten-spezifischen Kostenfunktion. Zum ersten Mal wurden unterschiedliche bildbasierte Kostenfunktionen in Bezug auf Segmentierungsgenauigkeit miteinander verglichen und die Parameterwahl wurde erstmalig durch ein datengetriebenes Verfahren durchgeführt. Unterschiedliche Lymphknotentypen, insbesondere pathologische und mit geringem Bildkontrast an der Grenze, werden deutlich besser unterstützt, als durch in der Literatur beschriebene Verfahren geschieht.

Darüber hinaus wurden auch zum ersten Mal unterschiedliche Optimierungsverfahren gegenübergestellt. Weitere Neuerungen umfassen eine Kostenfunktion, welche aufgrund kumulierter Bildintensitäten außerhalb eines vorgegebenen Intervalls gewisse Bildbereiche ausschließt, eine interaktive Erweiterung, welche es dem Anwender erlaubt in den Segmentierungsprozess einzugreifen, ein bildgradientenbasiertes Formwissen, welches den unterstützten Formbereich vergrößert, sowie eine Saatpunkt-invariante Erweiterung, die zu einer geringeren Inter- und Intra-Observer-Variabilität führt.

**Anwendungen:** Neben der Segmentierung von Lymphknoten eignet sich das Verfahren insbesondere für Wachstumsstrukturen mit hoher Form- und Größenvarianz, wie z.B. Tumore. Auch kann das Verfahren bei rundlichen Organen, wie z.B. dem Auge oder der Prostata, als Alternative zu modellbasierten Segmentierungsverfahren verwendet werden. Ein umfangreiches Training der Form entfällt.

### 7.2. Bildregistrierung auf Grundlage rigider Strukturen

Das in dieser Arbeit entwickelte Registrierungsverfahren wurde zur Intra-Subjekt-Registrierung von multimodalen/multitemporalen Kopf-Hals-Tomographien entwickelt und evaluiert. Zielorgane waren dabei die Lymphknoten im Kopf-Hals-Bereich.

**Schwierigkeiten:** Das Registrierungsverfahren soll sowohl unterschiedliche Bildmodalitäten als auch Aufnahmen unterschiedlicher Zeitpunkte unterstützen. Änderungen in der Anatomie sollen dabei weitestgehend erhalten bleiben. Im klinischen Alltag aufgenommene Aufnahmen weisen oftmals starke Bildartefakte auf, es werden unterschiedliche Bildausschnitte angezeigt und die Lage des Patienten variiert.

**Stand der Forschung:** Bildbasierte rigide Registrierungsverfahren existieren seit Anfang der 1990er Jahre und haben Einzug in den klinischen Alltag gefunden. Sie sind vollautomatisch, erhalten Anatomieänderungen und sind sehr robust. Allerdings unterstützen sie prinzipbedingt unterschiedliche Patientenlagen nicht. Bildbasierte deformierbare Verfahren können dies, besitzen allerdings einen hochdimensionalen Parameterraum und sind daher komplex und nicht robust gegenüber Bildartefakten und komplementärer Information unterschiedlicher Bildmodalitäten. Unterschiedliche Arten von lokaler Regularisierung vermindern diesen Effekt. Landmarkenbasierte deformierbare Registrierungsverfahren erfordern die robuste Segmentierung korrespondierender Landmarken, was insbesondere bei unterschiedlichen Bildmodalitäten eine große Herausforderung ist. Multi-rigide Verfahren, die auf Grundlage der rigiden Registrierung einzelner rigider Strukturen (Knochen) ein dichtes Deformationsfeld berechnen, bieten einen guten Kompromiss zwischen den beiden Extremen, erfordern allerdings meistens die Segmentierung der einzelnen rigiden Strukturen in mindestens einer der Bildmodalitäten. Daher existiert kein generalisierbares automatisches Registrierungsverfahren dieser Art, welches auch Aufnahmen unterschiedlicher Bildmodalitäten unterstützt. Ein Vergleich des in dieser Arbeit entwickelten Verfahrens mit existierenden Verfahren ist in Tabelle 7.2 zu finden.

**Prinzip:** Nach einer rigiden Registrierung des Kopfes wird ein artikulierter Atlas zur Segmentierung der einzelnen Knochen aus der CT-Aufnahme verwendet. Dieser Atlas basiert auf Wissen über Form und Aussehen einzelner Knochen und deren örtlichem Zusammenhang (Artikulation). Dieses Wissen wurde aus annotierten Trainings-CT-Aufnahmen gelernt. Eine sukzessiv lokal rigide Registrierung oder ein um personalisiertes Wissen angereicherter artikulierter Atlas bestimmt einzelne rigide Transformationen der Knochen zur anderen Aufnahme (andere Modalität oder Zeitserie). Unter der Annahme von Kontinuität wird dann aus diesen Transformationen ein dichtes Gesamttransformationsfeld durch ein zweistufiges Verfahren berechnet. Optional wird die aus beiden Aufnahmen extrahierte

| Verfahren           | Jahr | Intra-Subjekt | Voll automatisches Verfahren | Multi-Modal möglich | Multi-Rigide | Automatische Seg. benötigter Teile | Abhängige Reg. der rigiden Teile | Behandlung des Weichteilgewebes | Anwendung auf Kopf-Hals | Quant. Evaluation der Weichteile |
|---------------------|------|---------------|------------------------------|---------------------|--------------|------------------------------------|----------------------------------|---------------------------------|-------------------------|----------------------------------|
| Hill[97]            | 1994 | ✓             | ✗                            | ✓                   | ✗            | -                                  | -                                | -                               | ✓                       | ✓                                |
| Little[54]          | 1996 | ✓             | ✗                            | -                   | ✓            | ✗                                  | ✗                                | INT                             | ✓                       | ✗                                |
| Studholme[99]       | 1996 | ✓             | ✓                            | ✓                   | ✗            | -                                  | -                                | -                               | ✓                       | ✓                                |
| Huesmann[120]       | 2003 | ✓             | ✗                            | ✓                   | ✓            | ✗                                  | ✗                                | INT                             | ✗                       | ✗                                |
| Hu[121]             | 2004 | ✓             | ✗                            | ✓                   | ✓            | ✗                                  | ✗                                | -                               | ✗                       | ✗                                |
| Papademetris[126]   | 2005 | ✓             | ✗                            | ✗                   | ✓            | ✗                                  | ✓                                | INT                             | ✗                       | ✗                                |
| Cech[122]           | 2006 | ✓             | ✓                            | ✓                   | ✓            | ✓                                  | ✗                                | INT                             | ✗                       | ✗                                |
| Li[124]             | 2006 | ✓             | ✓                            | ✗                   | ✓            | -                                  | ✓                                | REG                             | ✗                       | ✗                                |
| Urschler[118]       | 2006 | ✓             | ✓                            | ✗                   | ✓            | ✓                                  | ✗                                | INT                             | ✓                       | ✗                                |
| Baiker[125]         | 2007 | ✓             | ✓                            | ✗                   | ✓            | ✓                                  | ✓                                | -                               | ✗                       | ✗                                |
| D.B. d'Aische[117]  | 2007 | ✓             | ✗                            | ✓                   | ✓            | ✗                                  | ✓                                | FEM                             | ✓                       | ✓                                |
| Commowick[68]       | 2008 | ✗             | ✗                            | ✗                   | ✗            | -                                  | -                                | REG                             | ✓                       | ✓                                |
| Han[112]            | 2008 | ✗             | ✓                            | ✗                   | ✗            | -                                  | -                                | REG                             | ✓                       | ✓                                |
| Klinder[133]        | 2008 | -             | ✓                            | -                   | ✓            | ✓                                  | ✓                                | -                               | ✗                       | -                                |
| Li[123]             | 2008 | ✓             | ✗                            | ✗                   | ✓            | ✗                                  | ✗                                | REG                             | ✗                       | ✗                                |
| Papastavrou [111]   | 2009 | ✓             | ✓                            | ✗                   | ✗            | -                                  | -                                | -                               | ✓                       | ✗                                |
| Gorthi[70]          | 2009 | ✗             | ✗                            | ✗                   | ✗            | -                                  | -                                | REG                             | ✓                       | ✓                                |
| M.-Fernández [128]  | 2009 | ✓             | ✓                            | ✗                   | ✓            | ✓                                  | ✓                                | INT                             | ✗                       | ✓                                |
| Al-Mayah[119]       | 2010 | ✓             | ✗                            | ✗                   | ✓            | ✗                                  | ✗                                | FEM                             | ✓                       | ✓                                |
| Gering[109]         | 2010 | ✗             | ✓                            | ✗                   | ✗            | -                                  | -                                | REG                             | ✓                       | ✓                                |
| Hollensen[106]      | 2010 | ✗             | ✓                            | ✗                   | ✗            | -                                  | -                                | REG                             | ✓                       | ✓                                |
| Ramus[107]          | 2010 | ✗             | ✓                            | ✗                   | ✗            | -                                  | -                                | REG                             | ✓                       | ✓                                |
| Teng[71]            | 2010 | ✗             | ✓                            | ✗                   | ✗            | -                                  | -                                | REG                             | ✓                       | ✓                                |
| Yang[110]           | 2010 | ✗             | ✓                            | ✗                   | ✗            | -                                  | -                                | REG                             | ✓                       | ✓                                |
| Suh[127]            | 2011 | ✓             | ✓                            | ✗                   | ✓            | -                                  | ✓                                | REG                             | ✗                       | ✓                                |
| <b>Diese Arbeit</b> | 2012 | ✓             | ✓                            | ✓                   | ✓            | ✓                                  | ✓                                | INT                             | ✓                       | ✓                                |

Tabelle 7.2.: Zusammenfassung von Eigenschaften des in dieser Arbeit entwickelten deformierbaren multi-rigiden Kopf-Hals-Registrierungsverfahrens mit dem Stand der Forschung (vgl. Tabelle 3.2). Nicht relevante/beschriebene Merkmale sind durch - gekennzeichnet. Die Behandlung des Weichteilgewebes kann entweder durch bildbasierte Registrierung (REG), Finite-Elemente-Methoden (FEM) oder Interpolation (INT) bestimmt werden.

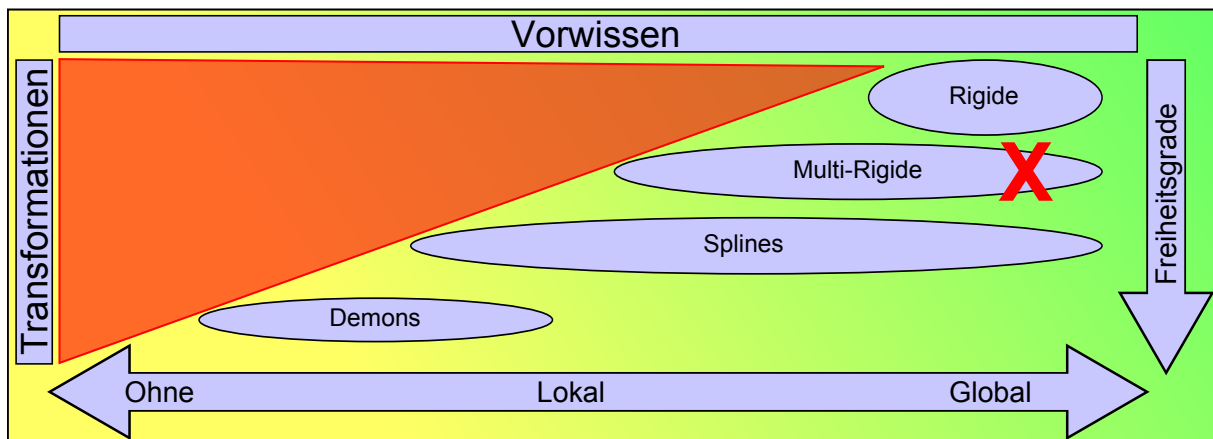


Abbildung 7.1.: Einordnung des Verfahrens zur Registrierung von Kopf-Hals-Aufnahmen (markiert durch ein X) in das in Kapitel 2.2.2.2 präsentierte Schema

Körperoberfläche als weiterer Stützpunkt in den Registrierungsprozess einbezogen. Um eine besonders hohe Registrierungsgenauigkeit in den Lymphknoten zu erreichen, wird dort abschließend ein lokal rigides Registrierungsverfahren angewendet.

Nach dem in Abschnitt 2.2.2 vorgestellten Klassifikationsschema kann sowohl das Registrierungsverfahren als auch der artikulierte Atlas als *Regmentierungsverfahren* betrachtet werden. Bei Ersterem wird ein Segmentierungsverfahren eingesetzt, welches einzelne rigide Strukturen identifiziert, um diese anschließend als Grundlage für ein Registrierungsverfahren zu verwenden. Bei Letzterem findet hingegen eine Registrierung zwischen dem Referenzbild und dem artikulierten Atlas statt, um daraus eine Segmentierung zu erhalten.

Weiterhin kann das multi-rigide Registrierungsverfahren in Bezug auf Transformationsklasse und Art des verwendeten Vorwissens kategorisiert werden. Wie in Abbildung 7.1 zu sehen ist, handelt es sich um ein multi-rigides Verfahren mit globaler Regularisierung. Für ein nicht-rigides Verfahren hat dies eine sehr geringe Anzahl an Freiheitsgraden. Dies hat zur Folge, dass zum einen die Komplexität des Verfahrens verhältnismäßig gering ist und zum anderen die Robustheit recht hoch ist, weil nur die Transformationen zulässig sind, die für diesen ganz bestimmten Anwendungsfall infrage kommen. Im Gegenzug erfordert dieses Verfahren das erneute Durchlaufen eines Trainingsprozesses, wenn es auf andere Registrierungsprobleme angewendet werden soll.

**Evaluation:** Innerhalb von Lymphknotenzentren wird eine Registrierungsgenauigkeit von durchschnittlich 5.05 mm erreicht. Diese ist tendenziell umso höher, je näher sich eine Struktur an den Knochen befindet und desto geringer die auszugleichenden Deformationen sind. Gegenüber dem Standardverfahren der nicht-rigiden bildbasierten Registrierung – der B-Spline Registrierung – ergibt sich eine Verbesserung der Registrierungsgenauigkeit von 37%. Eine anschließende lokal rigide Registrierung der Lymphknoten kann die Genauigkeit für bereits gut vorregistrierte Lymphknoten noch einmal verbessern, so dass bei 42% anstatt 27% der Lymphknotenzentren ein Registrierungsfehler von unter 3 mm erzielt wird. Die Qualität der erzeugten deformierten Bilder wird subjektiv als deutlich besser empfunden als dies bei bildbasierten Verfahren der Fall ist. Ein Grund dafür ist, dass Bildartefakte im Weichteilgewebe keinen Einfluss auf das Deformationsfeld haben und auch innerhalb der Kno-



chen durch die geringe Anzahl an Freiheitsgraden und die globale Regularisierung nur bedingt einen negativen Einfluss auf das Deformationsfeld haben. Das vorgestellte Verfahren benötigt in der Regel 1 – 2 Minuten zur kompletten Registrierung des Kopf-Hals-Bereichs. Die automatische Segmentierung der einzelnen rigiden Strukturen (Knochen) aus der CT-Aufnahme war der anspruchsvollste Teil des Verfahrens. Im Stand der Forschung ist dazu in der Regel Interaktion nötig. Der artikulierte Atlas, welcher sich einfach auch auf andere Bereiche des Körpers anwenden lässt, erzielt dazu eine Erfolgsrate von 86% bei einem Training auf gerade einmal 15 Datensätzen. Eine Vergrößerung der Trainingsmenge oder eine Initialisierung durch ein globales Objektdetektionsverfahren könnte dies noch einmal deutlich erhöhen.

**Eigene Beiträge:** In dieser Arbeit wurde das erste generalisierbare vollautomatische multi-rigide Registrierungsverfahren vorgestellt, welches auch unterschiedliche Bildmodalitäten unterstützt. Die Verwendung des (personalisierten) artikulierten Atlas verleiht dem Verfahren eine einmalige globale Regularisierung. Spezifische Beiträge dieser Arbeit sind:

- **Gelernte Artikulation in der atlasbasierten Segmentierung:** Der in dieser Arbeit vorgestellte artikulierte Atlas lernt die örtlichen Zusammenhänge mehrerer Strukturen aus Trainingsdaten. In der Literatur geschieht dies, wenn überhaupt, üblicherweise durch eine explizite Modellierung.
- **Adaptive Anpassung des artikulierten Atlas:** Bei der Anpassung des artikulierten Atlas ändert sich der Einfluss der Bildenergie für jedes Element im Laufe der Anpassung, wodurch eine lokale Anpassung an falsche Strukturen vermieden wird. In der Literatur erfolgt die Anpassung üblicherweise für alle Strukturen gleichzeitig.
- **Personalisierter artikulierter Atlas:** Die Personalisierung des artikulierten Atlas ermöglicht eine globale Regularisierung zur Bildregistrierung. Darüber hinaus ermöglicht es auch die erfolgreiche Segmentierung von sehr schwach kontrastierten Strukturen, wie z.B. Knochen in MRT-Aufnahmen.
- **Deformationsfeldberechnung:** Im Gegensatz zu den meisten interpolationsbasierten Deformationsfeldberechnungsmethoden ist die Linearkombination von Transformationen durch die Verwendung des Matrixexponentials mathematisch fundiert und auch Rotationsanteile werden propagiert. Während des zweistufigen Prozesses wird bedingt Wissen über die Deformation von unterschiedlichem Material (Gewebe/Luft) eingebracht.
- **Robuste rigide Registrierung:** Im Gegensatz zum Standardverfahren zur multimodalen rigiden Registrierung ist das in dieser Arbeit vorgestellte Verfahren durch die automatische Extraktion einer gemeinsamen Landmarke in beiden Aufnahmen in der Lage, sowohl mit starken Deformationen des Halses als auch mit stark beschnittenen Aufnahmen zurechtzukommen.

**Anwendungen:** Neben der Registrierung von Kopf-Hals-Tomographien, und damit auch der darin enthaltenen Lymphknoten, eignet sich das Verfahren für alle Körperbereiche, deren Deformation aufgrund rigider Strukturen zustande kommt. Dies sind z.B. der Brustkorb, die Hüfte und die Extremitäten. Selbst eine deformierbare Ganzkörperregistrierung ist denkbar. Um andere Bereiche registrieren zu können ist lediglich ein erneutes Training des artikulierten Atlas auf repräsentativen annotierten Aufnahmen nötig.

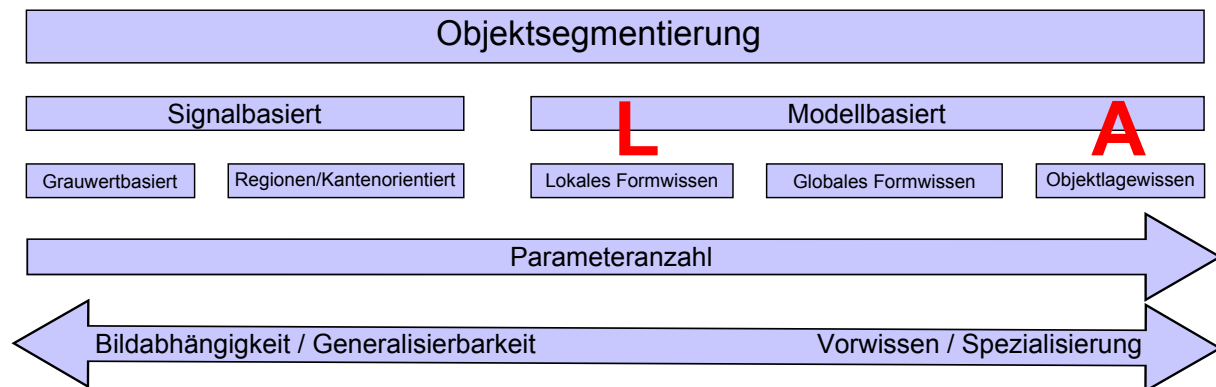


Abbildung 7.2.: Einordnung des in dieser Arbeit entwickelten Lymphknoten-Segmentierungsverfahrens (markiert durch ein L) und des artikulierte Atlas zur Segmentierung einzelner Knochen (markiert durch ein A) in das in Kapitel 2.2.1 vorgestellte Klassifikationsschema

### 7.3. Konzeptioneller Vergleich der Radialstrahl-basierten Segmentierung mit dem artikulierte Atlas

Im Rahmen dieser Arbeit wurden zwei fundamental unterschiedliche Segmentierungsverfahren entwickelt, die nun miteinander konzeptionell verglichen werden. Dies ist zum einen das Radialstrahl-basierte Segmentierungsverfahren mit lokalem Formwissen und zum anderen der lageorientierte artikulierte Atlas zur gleichzeitigen Segmentierung mehrerer Knochen (siehe Abbildung 7.2). Bei ersterem ist wenig über die Struktur – z.B. den Lymphknoten – bekannt. Größe sowie Form variieren stark und die Lage ist relativ zu anderen Strukturen beinahe beliebig. Daher kann nur wenig Vorwissen verwendet werden und das Signal – also das Bild – ist die maßgebliche Segmentierungsgrundlage. Da jedoch in vielen Fällen der Kontrast zu benachbarten Strukturen sehr gering ist, kann auf Vorwissen nicht ganz verzichtet werden. Lokales Vorwissen über die Form stellt einen Kompromiss zwischen den widersprüchlichen Anforderungen dar.

Bei der Segmentierung der Knochen und insbesondere der Wirbel stellt sich die Situation grundsätzlich verschieden dar. Die Varianz an Form und Größe ist vergleichsweise gering, es kann also an dieser Stelle viel Vorwissen eingebracht werden. Dies allein ist allerdings nicht ausreichend, weil benachbarte Wirbel sehr ähnlich aussehen und so die bloße Berücksichtigung der Form eines Wirbels schier unmöglich eine Unterscheidung benachbarter Wirbel zulässt. Daher muss zusätzlich auf Wissen über die relative Lage benachbarter Strukturen während des Segmentierungsprozesses zurückgegriffen werden. Dies geschieht durch den artikulierte Atlas. Die Kehrseite ist, dass das Verfahren sehr speziell auf einen ganz bestimmten Anwendungsfall – in diesem Fall die Segmentierung der Knochen aus Kopf-Hals-Aufnahmen – zugeschnitten ist und eine andere Anwendung ein erneutes Training erfordert.

## 7.4. Weiterführende Arbeit

Während der Beantwortung der Fragestellung dieser Arbeit eröffneten sich weitere Fragestellungen, die sich nicht im Rahmen dieser Arbeit befinden, allerdings die Ergebnisse dieser Arbeit aufgreifen und damit weitere Anwendungsszenarien erschließen. Im Folgenden sind die wichtigsten aufgelistet:

**Detektion von Unterschieden in Zeitserien CT-Aufnahmen:** Im Gegensatz zu allen bildbasierten Registrierungsverfahren, zeichnet sich das in dieser Arbeit vorgestellte Registrierungsverfahren dadurch aus, dass es nur Unterschiede aufgrund verschiedener Patientenlage ausgleicht, nicht jedoch aufgrund einer Änderung der Anatomie. Im Fall von Zeitserien-CT-Aufnahmen kann dies dazu genutzt werden, um Anatomieänderungen unterschiede zu visualisieren oder wenn man einen Schritt weiter geht, automatisch zu detektieren. Unter der Annahme, dass es sich nach der Registrierung mit dem in dieser Arbeit vorgestellten Verfahren bei allen vorhandenen Bildunterschieden um Anatomieunterschiede handelt, kann eine Analyse des erhaltenen Deformationsfeldes einer nachgeschalteten bildbasierten deformierbaren Registrierung dies erreichen. Eine Visualisierung der Deformationsstärke hebt Orte mit starken Anatomieunterschieden hervor. Klassifikationsverfahren können verwendet werden, um Anatomieunterschiede daraus automatisch zu detektieren.

**Erhöhung der Registrierungsgenauigkeit im Weichteilgewebe:** Zur weiteren Erhöhung des in dieser Arbeit vorgestellten Registrierungsverfahrens, sind mehrere Möglichkeiten denkbar, die es wert sind zu verfolgen:

- Die Einbeziehung der Körperoberfläche konnte die Registrierungsgenauigkeit erhöhen, allerdings stieg die Anfälligkeit für Bildartefakte und die deformierten Bilder sahen teilweise unnatürlich aus. Eine Verbesserung der Körperoberflächenextraktion könnte letzteres verhindern. Modellbasierte Segmentierungsverfahren scheinen dafür geeignet zu sein.
- Weitere Stützstrukturen aus dem Weichteilgewebe könnten in den Registrierungsprozess einbezogen werden. Idealerweise sollten diese in allen betrachteten Bildmodalitäten einfach zu segmentieren sein. Ein Beispiel dafür ist die Trachea.
- Die Registrierungsgenauigkeit war insbesondere im unteren Halsbereich schlechter als im oberen Halsbereich und dem Kopf. Der Grund dafür ist hauptsächlich, dass bisher nur die Knochen der Halswirbelsäule und die ersten beiden Brustwirbel als Stützstellen in den Registrierungsprozess eingebunden wurden. Tatsächlich wirken sich aber auch die Transformationen anderer Knochen auf die Registrierungsgenauigkeit aus und sollten ebenfalls einbezogen werden. Dies sind neben den Knochen im Schulterbereich auch die obersten Rippen. Problematisch war im Rahmen der Fragestellung hier dabei jedoch, dass nicht alle dieser Knochen in allen Aufnahmen dargestellt sind und somit beim Lernen des Artikulationsraums das Fehlen von Daten geeignet berücksichtigt werden muss oder bei der Datenakquise darauf im Vorfeld geachtet werden muss.
- Das bisher verwendete Modell zur Propagierung der Transformationen von den Knochen in das umliegende Weichteilgewebe fordert lediglich, dass die Transformationen benachbarter Orte ähnlich sind. Die Verwendung eines biomechanischen Modells, welches bspw. Sehnen und die damit verbundenen Restriktionen der Bewegung in den Modellierungsprozess einbezieht oder unterschiedlichen Gewebetypen unterschiedliche Eigenschaften zuordnet, könnte die Registrie-

rungsgenauigkeit verbessern. Jedoch ist damit die Segmentierung aller Strukturen erforderlich, die in das biomechanische Modell einbezogen werden sollen.

**Kompensation der Atembewegung:** Wird der Brustkorb durch den artikulierten Atlas modelliert, kann dieser auch dahingehend erweitert werden, dass eine Kompensation der Atembewegung möglich ist. Dazu müssten beim Lernen des Artikulationsraums von mehreren Patienten annotierte CT-Aufnahmen von verschiedenen Zuständen des gesamten Atemzyklus vorliegen. Die Atembewegung kann dann getrennt von der Anatomievarianz gelernt werden. Wird der artikulierte Atlas des Brustkorbs dann an einen ungesesehenen Datensatz angepasst, kann daraus dann der aktuelle Atemzustand bestimmt werden und die CT-Aufnahme in einen beliebigen Atemzustand deformiert werden.

**Artikulierter Atlas zur Segmentierung:** Bisher wurde der artikulierte Atlas nur zur Segmentierung einzelner Knochen verwendet. Dabei könnte dieser auch auf andere Strukturen erweitert werden. Im Kopf-Hals-Bereich ist dies vor allem zur Strahlentherapieplanung interessant, da hier Zielorgane und eine Vielzahl von Strukturen segmentiert werden müssen. Dies sind z.B. der Hirnstamm, der Spinalkanal, Speicheldrüsen, die Augen, der Sehnerv und die Lymphknotenregionen. Da die Strukturen voneinander abhängen, bietet sich die gemeinsame Segmentierung durch ein lageorientiertes Verfahren, wie dem artikulierten Atlas, an. Die meisten Strukturen weisen allerdings keinen so hohen Kontrast wie Knochen auf und auch die Formvarianz ist deutlich höher. Ersters kann ideal durch Wissen über die relative Lage von Strukturen mit schwachem Kontrast zu Strukturen mit starkem Kontrast berücksichtigt werden. Letzteres kann durch die Kombination von statistischen Formmodellen mit dem artikulierten Atlas abgedeckt werden.

**Detektion von Lymphknoten:** Die Segmentierung von Lymphknoten durch das in dieser Arbeit vorgestellte Verfahren erfordert die Initialisierung durch einen Saatpunkt in der Nähe des Objektzentrums. Für diese Arbeit ist dies erwünscht, weil es dem Anwender auf einfache Art und Weise ermöglicht, den zu untersuchenden Lymphknoten auszuwählen. Soll dies allerdings automatisiert werden, ist ein automatisches Verfahren zur Detektion von Lymphknoten erforderlich. Einige Arbeiten dazu existieren bereits [64, 65, 66, 67], jedoch ist die hohe Falsch-Positiv-Rate nach wie vor ein Problem. Durch die Segmentierung einer Vielzahl von Strukturen durch den artikulierten Atlas können sehr viele Bereiche der Aufnahme ausgeschlossen werden. Die Detektion könnte sich dann auf die designierte Lymphknotenregionen beschränken und so das Detektionsergebnis vor allem in puncto Falsch-Positiv-Rate verbessern.

# Eigene Literatur

- [1] Sebastian Steger, Fabio Franco, Nicola Sverzellati, Gianfranco Chiari, and Ramon Colomer, “3d assessment of lymph nodes vs. recist 1.1”, *Academic Radiology*, vol. 18, no. 3, pp. 391 – 394, 2011.
- [2] Sebastian Steger, Marius Erdt, Gianfranco Chiari, and Georgios Sakas, “Feature extraction from medical images for an oral cancer reoccurrence prediction environment”, in *World Congress on Medical Physics and Biomedical Engineering, September 7 - 12, 2009, Munich, Germany*, Olaf Dössel, Wolfgang C. Schlegel, and Ratko Magjarevic, Eds., vol. 25/5 of *IFMBE Proceedings*, pp. 97–100. Springer Berlin Heidelberg, 2009.
- [3] Marco Picone, Sebastian Steger, Konstantinos Exarchos, Marco Fazio, Yorgos Goletsis, Dimitrios I. Fotiadis, Elena Martinelli, and Diego Ardigó, “Enabling heterogeneous data integration and biomedical event prediction through ict: The test case of cancer reoccurrence”, in *Software Tools and Algorithms for Biological Systems*, Hamid R. R. Arabnia and Quoc-Nam Tran, Eds., vol. 696 of *Advances in Experimental Medicine and Biology*, pp. 367–375. Springer New York, 2011.
- [4] Dario Salvi, Marco Picone, Maria Theresa Arredondo, Maria Fernanda Cabrera-Umpierrez, Angel Esteban, Sebastian Steger, and Tito Poli, “Merging person-specific bio-markers for predicting oral cancer recurrence through an ontology”, *Biomedical Engineering, IEEE Transactions on*, vol. 60, no. 1, pp. 216–220, 2013.
- [5] Marius Erdt, Sebastian Steger, and Georgios Sakas, “Regmentation: A new view of image segmentation and registration”, *Journal of Radiation Oncology Informatics*, vol. 4, no. 1, 2012.
- [6] Marius Erdt, Sebastian Steger, Matthias Kirschner, and Stefan Wesarg, “Fast automatic liver segmentation combining learned shape priors with observed shape deviation”, in *Computer-Based Medical Systems (CBMS), 2010 IEEE 23rd International Symposium on*. IEEE, 2010, pp. 249–254.
- [7] Marius Erdt, Sebastian Steger, and Stefan Wesarg, “Deformable registration of mr images using a hierarchical patch based approach with a normalized metric quality measure”, in *Proc. IEEE Int Biomedical Imaging: From Nano to Macro Symp*, 2012.
- [8] Sebastian Steger and Marius Erdt, “Lymph node segmentation in ct images using a size invariant mass spring model”, in *Proc. 10th IEEE Int Information Technology and Applications in Biomedicine (ITAB) Conf*, 2010, pp. 1–4.
- [9] Sebastian Steger, Daniel Ebert, and Marius Erdt, “Lymph node segmentation in ct slices using dynamic programming”, in *Proc. IEEE Int Biomedical Imaging: From Nano to Macro Symp*, 2011, pp. 1990–1993.
- [10] Sebastian Steger, Nazli Bozoglu, Arjan Kuijper, and Stefan Wesarg, “Application of radial ray

- based segmentation to cervical lymph nodes in ct images", *Medical Imaging, IEEE Transactions on*, 2013.
- [11] Sebastian Steger and Georgios Sakas, "Image gradient based shape prior for segmentation of not that spherical strcutures", in *Proc. IEEE Int Biomedical Imaging: From Nano to Macro Symp*, 2012.
- [12] Sebastian Steger and Georgios Sakas, *Abdominal Imaging: Computational and Clinical Applications*, vol. 7029 of *Lecture Notes in Computer Science*, chapter FIST: Fast Interactive Segmentation of Tumors, Springer, March 2012.
- [13] Sebastian Steger and Stefan Wesarg, "Automated skeleton based multi-modal deformable registration of head&neck datasets", in *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2012*, Nicholas Ayache, Hervé Delingette, Polina Golland, and Kensaku Mori, Eds., vol. 7511 of *Lecture Notes in Computer Science*, pp. 66–73. Springer Berlin / Heidelberg, 2012.
- [14] Sebastian Steger and Matthias Keil, "Automated initialization and region of interest detection for successful head registration of truncated ct/mr head & neck images", in *Proc. 10th IEEE Int Information Technology and Applications in Biomedicine (ITAB) Conf*, 2010, pp. 1–5.
- [15] Sebastian Steger, Matthias Kirschner, and Stefan Wesarg, "Articulated atlas for segmentation of the skeleton from head & neck ct datasets", in *Proc. IEEE Int Biomedical Imaging: From Nano to Macro Symp*, 2012.
- [16] Sebastian Steger, "Local rigid registration for multimodal texture feature extraction from medical images", in *Medical Imaging 2011: Image Processing*, Benoit M. Dawant and David R. Haynor, Eds. 2011, vol. 7962, p. 79621V, SPIE.

## Betreute Abschlussarbeiten

- [17] Nazli Bozoglu, "Evaluierung von algorithmen zur segmentierung von lymphknoten in ct-datensätzen", Master's thesis, Technische Universität Darmstadt, 2012, Betreuer: Sebastian Steger.
- [18] Dmytro Vronskyi, "Rigide registrierung von 3d anatomischen mri und 2d biochemischen mrsi daten", Master's thesis, Technische Universität Darmstadt, 2010, Betreuer: Sebastian Steger.





# Literaturverzeichnis

- [19] Jacques Ferlay, Hai-Rim Shin, Freddie Bray, David Forman, Colin Mathers, and Donald Maxwell Parkin, "Estimates of worldwide burden of cancer in 2008: Globocan 2008", *International Journal of Cancer*, vol. 127, no. 12, pp. 2893–2917, 2010. 1
- [20] P. Boyle and B. Levin, Eds., *World Cancer Report 2008*, World Health Organization, 2009. 1
- [21] Jeffrey Myers, Ed., *Oral Cancer Mestastasis*, Springer, 2010. 1, 2
- [22] Robert Hermans, Ed., *Head and Neck Cancer Imaging*, Springer, 2006. 1, 2
- [23] L. H. Sobin, Mary K. Gospodarowicz, and Christian Wittekind, Eds., *TNM Classification of Malignant Tumours*, Uicc International Union Against Cancer, 2009. 2
- [24] P Therasse, S G Arbuck, E A Eisenhauer, J Wanders, R S Kaplan, L Rubinstein, J Verweij, M Van Glabbeke, A T Van Oosterom, M C Christian, and et al., "New guidelines to evaluate the response to treatment in solid tumors. european organization for research and treatment of cancer, national cancer institute of the united states, national cancer institute of canada.", *Journal Of The National Cancer Institute*, vol. 92, no. 3, pp. 205–216, 2000. 2
- [25] E.A. Eisenhauer, P. Therasse, J. Bogaerts, L.H. Schwartz, D. Sargent, R. Ford, J. Dancey, S. Arbuck, S. Gwyther, M. Mooney, L. Rubinstein, L. Shankar, L. Dodd, R. Kaplan, D. La-combe, and J. Verweij, "New response evaluation criteria in solid tumours: Revised recist guideline (version 1.1)", *European Journal of Cancer*, vol. 45, no. 2, pp. 228–247, Jan. 2009. 2
- [26] L.H. Schwartz, J. Bogaerts, R. Ford, L. Shankar, P. Therasse, S. Gwyther, and E.A. Eisenhauer, "Evaluation of lymph nodes with recist 1.1", *European Journal of Cancer*, vol. 45, no. 2, pp. 261–267, Jan. 2009. 2
- [27] Jonas Castelijns and Michiel van den Brekel, "Imaging of lymphadenopathy in the neck", *European Radiology*, vol. 12, pp. 727–738, 2002.
- [28] K. Thomas Robbins, Garry Clayman, Paul A. Levine, Jesus Medina, Roy Sessions, Ashok Shaha, Peter Som, Gregory T. Wolf, , the Committee for Head, Neck Surgery, American Academy of Otolaryngology-Head Oncology, and Neck Surgery, "Neck dissection classification update: Revisions proposed by the american head and neck society and the american academy of otolaryngology-head and neck surgery", *Arch Otolaryngol Head Neck Surg*, vol. 128, no. 7, pp. 751–758, 2002.
- [29] Par H. Rouvière, "Anatomie des lymphatiques de l'homme.", *Journal of the American Medical Association*, vol. 99, no. 20, pp. 1716, 1932.
- [30] T.H. Schiebler and H. Korf, *Anatomie: Histologie, Entwicklungsgeschichte, Makroskopische Und Mikroskopische Anatomie, Topographie*, Steinkopff-Verlag Darmstadt, 2007.

- [31] Takashi Nakamura and Misa Sumi, "Nodal imaging in the neck: recent advances in us, ct and mr imaging of metastatic nodes", *European Radiology*, vol. 17, pp. 1235–1241, 2007.
- [32] Martin G. Mack, Jörg Rieger, Mehran Baghi, Sotirios Bisdas, and Thomas J. Vogl, "Cervical lymph nodes", *European Journal of Radiology*, vol. 66, no. 3, pp. 493–500, 2008.
- [33] Heinz Handels, *Medizinische Bildverarbeitung*, Vieweg+Teubner Verlag / GWV Fachverlage GmbH, Wiesbaden, 2009.
- [34] J. S. Duncan and N. Ayache, "Medical image analysis: progress over two decades and the challenges ahead", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 1, pp. 85–106, 2000.
- [35] Nobuyuki Otsu, "A threshold selection method from gray-level histograms", *IEEE Transactions on Systems, Man and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.
- [36] S. Horowitz and T. Pavlidis, "Picture segmentation by a directed split-and-merge procedure", in *Proceedings Second International Joint Conference on Pattern Recognition*, Aug. 1974, pp. 424–433.
- [37] S. Beucher and F. Meyer, "The morphological approach to segmentation: the watershed transformation. mathematical morphology in image processing.", *Optical Engineering*, vol. 34, pp. 433–481, 1993.
- [38] John Canny, "A computational approach to edge detection", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, , no. 6, pp. 679–698, 1986.
- [39] Michael Kass, Andrew Witkin, and Demetri Terzopoulos, "Snakes: Active contour models", *International Journal of Computer Vision*, vol. 1, no. 4, pp. 321–331, 1988.
- [40] R. Malladi, J. A. Sethian, and B. C. Vemuri, "Shape modeling with front propagation: a level set approach", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, no. 2, pp. 158–175, 1995.
- [41] D. Greig, B. Porteous, and A. Seheult, "Exact maximum a posteriori estimation for binary images", *Royal Journal on Statistical Society*, vol. 51, no. 2, pp. 271–279, 1989.
- [42] Judea Pearl, *Probabilistic reasoning in intelligent systems: networks of plausible inference*, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1988.
- [43] K. P. Murphy, Y. Weiss, and M. Jordan, "Loopy belief propagation for approximate inference: An empirical study", in *Uncertainty in Artificial Intelligence*, 1999, pp. 467–475.
- [44] D Terzopoulos and K Fleischer, "Deformable models", *The Visual Computer*, vol. 4, no. 6, pp. 306–331, 1988.
- [45] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models and their training and application", *Computer Vision and Image Understanding*, vol. 61, pp. 38–59, January 1995.
- [46] Timothy F. Cootes, Gareth J. Edwards, and Christopher J. Taylor, "Active appearance models.", in *ECCV (2)*, Hans Burkhardt and Bernd Neumann, Eds. 1998, vol. 1407 of *Lecture Notes in Computer Science*, pp. 484–498, Springer.
- [47] Lars Dornheim, Klaus D. Tönnies, and Jana Dornheim, "Stable dynamic 3d shape models",

- in *In ICIP'05: International Conference on Image Processing*, 2005.
- [48] R O Duda and P E Hart, "Use of the hough transformation to detect lines and curves in pictures.", *Communications of the ACM*, vol. 15, no. 1, pp. 11–15, 1972.
- [49] A. Tsai, William Wells, Clare Tempany, E. Grimson, and Alan Willsky, "Coupled multi-shape model and mutual information for medical image segmentation", in *Information Processing in Medical Imaging*, Chris Taylor and J. Noble, Eds., vol. 2732 of *Lecture Notes in Computer Science*, pp. 185–197. Springer Berlin / Heidelberg, 2003.
- [50] D Louis Collins and A C Evans, "Animal: Validation and applications of nonlinear registration-based segmentation", *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 11, no. 8, pp. 1271–1294, 1997.
- [51] Fahmi Khalifa, Garth M. Beache, Georgy Gimel'farb, Jasjit S. Suri, and Ayman S. El-Baz, "State-of-the-art medical image registration methodologies: A survey", in *Multi Modality State-of-the-Art Medical Image Segmentation and Registration Methodologies*, Ayman S. El-Baz, Rajendra Acharya U, Majid Mirmehdi, and Jasjit S. Suri, Eds., pp. 235–280. Springer US, 2011.
- [52] J.B.Antoine Maintz and Max A. Viergever, "A survey of medical image registration", *Medical Image Analysis*, vol. 2, no. 1, pp. 1 – 36, 1998.
- [53] Vincent Arsigny, Olivier Commowick, Nicholas Ayache, and Xavier Pennec, "A fast and log-euclidean polyaffine framework for locally linear registration", *Journal of Mathematical Imaging and Vision*, vol. 33, pp. 222–238, 2009, 10.1007/s10851-008-0135-9.
- [54] J. A. Little, D. L. G. Hill, and D. J. Hawkes, "Deformations incorporating rigid structures", in *Proceedings on Workshop Mathematical Methods in Biomedical Image Analysis*, 1996, pp. 104–113.
- [55] M. H. Davis, A. Khotanzad, D. P. Flamig, and S. E. Harms, "A physics-based coordinate transformation for 3-d image matching", *IEEE Transactions on Medical Imaging*, vol. 16, no. 3, pp. 317–328, 1997.
- [56] Fred L. Bookstein, *Brain Warping*, chapter Linear Methods for Nonlinear Maps: Procrustes Fits, Thin-Plate Splines, and the Biometric Analysis of Shape Variability, pp. 157–181, Academic Press, 1999.
- [57] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. G. Hill, M. O. Leach, and D. J. Hawkes, "Nonrigid registration using free-form deformations: application to breast mr images", *IEEE Transactions on Medical Imaging*, vol. 18, no. 8, pp. 712–721, 1999.
- [58] J.-P. Thirion, "Image matching as a diffusion process: an analogy with maxwell's demons", *Medical Image Analysis*, vol. 2, no. 3, pp. 243 – 260, 1998.
- [59] G. Wahba, *Spline models for observational data*, vol. 59 of *CBMS-NSF Regional Conference Series in Applied Mathematics*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1990.
- [60] T. Rohlfing, Jr. Maurer, C. R., D. A. Bluemke, and M. A. Jacobs, "Volume-preserving non-rigid registration of mr breast images using free-form deformation with an incompressibility constraint", *IEEE Transactions on Medical Imaging*, vol. 22, no. 6, pp. 730–741, 2003.

- [61] Dirk Loeckx, Frederik Maes, Dirk Vandermeulen, and Paul Suetens, "Nonrigid image registration using free-form deformations with a local rigidity constraint", in *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2004*, Christian Barillot, David Haynor, and Pierre Hellier, Eds., vol. 3216 of *Lecture Notes in Computer Science*, pp. 639–646. Springer Berlin / Heidelberg, 2004.
- [62] C. Tanner, J. Schnabel, D. Chung, M. Clarkson, D. Rueckert, D. Hill, and D. Hawkes, "Volume and shape preservation of enhancing lesions when applying non-rigid registration to a time series of contrast enhancing mr breast images", in *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2000*, Scott Delp, Anthony DiGoia, and Branislav Jaramaz, Eds., vol. 1935 of *Lecture Notes in Computer Science*, pp. CH371–CH371. Springer Berlin / Heidelberg, 2000.
- [63] Kexiang Wang, Ying He, and Hong Qin, "Incorporating rigid structures in non-rigid registration using triangular b-splines", in *Variational, Geometric, and Level Set Methods in Computer Vision*, Nikos Paragios, Olivier Faugeras, Tony Chan, and Christoph Schnörr, Eds., vol. 3752 of *Lecture Notes in Computer Science*, pp. 235–246. Springer Berlin / Heidelberg, 2005.
- [64] Takayuki Kitasaka, Yukihiro Tsujimura, Yoshihiko Nakamura, Kensaku Mori, Yasuhito Suenaga, Masaaki Ito, and Shigeru Nawano, "Automated extraction of lymph nodes from 3-d abdominal ct images using 3-d minimum directional difference filter", in *Proceedings of the 10th international conference on Medical image computing and computer-assisted intervention*, Berlin, Heidelberg, 2007, MICCAI'07, pp. 336–343, Springer-Verlag.
- [65] Kongkuo Lu, Scott A. Merritt, and William E. Higgins, "Extraction and visualization of the central chest lymph-node stations", 2008, vol. 6915, p. 69151B, SPIE.
- [66] Lars Dornheim and Jana Dornheim, "Automatische detektion von lymphknoten in ct-datensätzen des halses.", in *Bildverarbeitung für die Medizin*, Thomas Tolxdorff, Jürgen Braun, Thomas Martin Deserno, Heinz Handels, Alexander Horsch, and Hans-Peter Meinzer, Eds. 2008, Informatik Aktuell, pp. 308–312, Springer.
- [67] A. Barbu, M. Suehling, X. Xu, D. Liu, S. K. Zhou, and D. Comaniciu, "Automatic detection and segmentation of lymph nodes from ct data", *IEEE Transactions on Medical Imaging*, vol. 31, no. 2, pp. 240–250, 2012.
- [68] Olivier Commowick, Vincent Grégoire, and Grégoire Malandain, "Atlas-based delineation of lymph node levels in head and neck computed tomography images", *Radiotherapy Oncology*, vol. 87, no. 2, pp. 281–289, May 2008.
- [69] Simon K Warfield, Kelly H Zou, and William M Wells, "Simultaneous truth and performance level estimation (staple): an algorithm for the validation of image segmentation.", *IEEE Transactions on Medical Imaging*, vol. 23, no. 7, pp. 903–21, 2004.
- [70] Subrahmanyam Gorthi, Valérie Duay, Nawal Houhou, Meritxell Bach Cuadra, Ulrike Schick, Minerva Becker, Abdelkarim S. Allal, and Jean-Philippe Thiran, "Segmentation of head and neck lymph node regions for radiotherapy planning, using active contour based atlas registration", *IEEE Journal of Selected Topics in Signal Processing*, vol. 3, no. 1, pp. 135–147, 2009.
- [71] Chia-Chi Teng, Linda G Shapiro, and Ira J Kalet, "Head and neck lymph node region delineation

- with image registration.", *Biomed Eng Online*, vol. 9, no. 30, pp. 30, 2010.
- [72] P. Viola and W. M. Wells III, "Alignment by maximization of mutual information", in *Proceedings on Conference Fifth International Computer Vision*, 1995, pp. 16–23.
- [73] A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Suetens, and G. Marchal, *Automated multi-modality image registration based on information theory*, vol. 14, pp. 263–274, Kluwer Academic Publishers, 1995.
- [74] Jadwiga Rogowska, Keith Barchelder, G. Scott Gazelle, Elkan F. Halpern, William Connor, and Gerald L. Wolf, "Evaluation of selected two-dimensional segmentation techniques for computed tomography quantitation of lymph nodes", *Investigative Radiology*, vol. 31, no. 3, pp. –, 1996.
- [75] O. A. Debats, G. J. S. Litjens, J. O. Barentsz, N. Karssemeijer, and H. J. Huisman, "Automated 3-dimensional segmentation of pelvic lymph nodes in magnetic resonance images", *Medical Physics*, vol. 38, no. 11, pp. 6178–6187, 2011.
- [76] G. Unal, G. Slabaugh, A. Ess, A. Yezzi, T. Fang, J. Tyan, M. Requardt, R. Krieg, R. Seethamraju, M. Harisinghani, and R. Weissleder, "Semi-automatic lymph node segmentation in In-mri", in *Proc. IEEE Int Image Processing Conference*, 2006, pp. 77–80.
- [77] Junhua Zhang, Yuanyuan Wang, and Xinling Shi, "An improved graph cut segmentation method for cervical lymph nodes on sonograms and its relationship with node's shape assessment.", *Computerized medical imaging and graphics the official journal of the Computerized Medical Imaging Society*, vol. 33, no. 8, pp. 602–607, 2009.
- [78] J. H. Moltz, L. Bornemann, J.-M. Kuhnigk, V. Dicken, E. Peitgen, S. Meier, H. Bolte, M. Fabel, H.-C. Bauknecht, M. Hittinger, A. Kiessling, M. Pusken, and H.-O. Peitgen, "Advanced segmentation techniques for lung nodules, liver metastases, and enlarged lymph nodes in ct scans", *IEEE Journal of Selected Topics in Signal Processing Inaugural Issue*, vol. 3, no. 1, pp. 122–134, 2009.
- [79] Daniel Maleike, Michael Fabel, Ralf Tetzlaff, Hendrik von Tengg-Koblighk, Tobias Heimann, Hans-Peter Meinzer, and Ivo Wolf, "Lymph node segmentation on ct images by a shape model guided deformable surface method", 2008, vol. 6914, p. 69141S, SPIE.
- [80] Jana Dornheim, Heiko Seim, Bernhard Preim, Ilka Hertel, and Gero Strauss, "Segmentation of neck lymph nodes in ct datasets with stable 3d mass-spring models: Segmentation of neck lymph nodes", *Academic Radiology*, vol. 14, no. 11, pp. 1389–1399, Nov. 2007.
- [81] Jana Dornheim, Heiko Seim, Bernhard Preim, Ilka Hertel, and Gero Strauss, "Segmentation of neck lymph nodes in ct datasets with stable 3d mass-spring models.", *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2006*, vol. 9, no. Pt 2, pp. 904–911, 2006.
- [82] Lars Dornheim, Jana Dornheim, and Ivo Rössling, "Complete fully automatic model-based segmentation of normal and pathological lymph nodes in ct data", *International Journal of Computer Assisted Radiology and Surgery*, vol. 5, pp. 565–581, 2010.
- [83] Lars Dornheim, Jana Dornheim, Ivo Rössling, and Tobias Mönch, "Model-based segmentation of pathological lymph nodes in ct data", 2010, vol. 7623, p. 76234V, SPIE.

- [84] L. Dornheim, "Quality of fit of stable mass-spring models", *Pattern Recognition and Image Analysis*, vol. 18, pp. 359–364, 2008.
- [85] Kongkuo Lu and William E. Higgins, "Segmentation of the central-chest lymph nodes in 3d mdct images", *Computers in Biology and Medicine*, vol. 41, no. 9, pp. 780 – 789, 2011.
- [86] Eric N. Mortensen and William A. Barrett, "Intelligent scissors for image composition", in *Proceedings of SIGGRAPH 95*, Aug. 1995, Computer Graphics Proceedings, Annual Conference Series, pp. 191–198.
- [87] David M. Honea, Yaorong Ge, Wesley E. Snyder, Paul F. Hemler, and David J. Vining, "Lymph node segmentation using active contours", 1997, vol. 3034, pp. 265–273, SPIE.
- [88] David M. Honea and Wesley E. Snyder, "Three-dimensional active surface approach to lymph node segmentation", 1999, vol. 3661, pp. 1003–1011, SPIE.
- [89] Jiayong Yan, Tian-ge Zhuang, Binsheng Zhao, and Lawrence H. Schwartz, "Lymph node segmentation from ct images using fast marching method", *Computerized Medical Imaging and Graphics*, vol. 28, no. 1, pp. 33–38, Jan. 2004.
- [90] Johannes Feulner, S. Kevin Zhou, Matthias Hammon, Joachim Hornegger, and Dorin Comaniciu, "Lymph node detection and segmentation in chest ct data using discriminative learning and a spatial prior", *Medical Image Analysis*, , no. 0, pp. –, 2012.
- [91] Jan Egger, Miriam Bauer, Daniela Kuhnt, Barbara Carl, Christoph Kappus, Bernd Freisleben, and Christopher Nimsky, "Nugget-cut: A segmentation scheme for spherically- and elliptically-shaped 3d objects", in *Pattern Recognition*, Michael Goesele, Stefan Roth, Arjan Kuijper, Bernt Schiele, and Konrad Schindler, Eds., vol. 6376 of *Lecture Notes in Computer Science*, pp. 373–382. Springer Berlin / Heidelberg, 2010.
- [92] Adrian Barbu, Michael Sühling, Xun Xu, David Liu, Shaohua Kevin Zhou, and Dorin Comaniciu, "Automatic detection and segmentation of axillary lymph nodes.", in *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2010*, Tianzi Jiang, Nassir Navab, Josien P. W. Pluim, and Max A. Viergever, Eds. 2010, vol. 6361 of *Lecture Notes in Computer Science*, pp. 28–36, Springer.
- [93] Yao Wang and Reinhard Beichel, "Graph-based segmentation of lymph nodes in ct data", in *Advances in Visual Computing*, George Bebis, Richard Boyle, Bahram Parvin, Darko Koracin, Ronald Chung, Riad Hammound, Muhammad Hussain, Tan Kar-Han, Roger Crawfis, Daniel Thalmann, David Kao, and Lisa Avila, Eds., vol. 6454 of *Lecture Notes in Computer Science*, pp. 312–321. Springer Berlin / Heidelberg, 2010.
- [94] Kang Li, Xiaodong Wu, Danny Z Chen, and Milan Sonka, "Optimal surface segmentation in volumetric images—a graph-theoretic approach.", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 1, pp. 119–134, 2006.
- [95] Reinhard R. Beichel and Yao Wang, "Computer-aided lymph node segmentation in volumetric ct data", *Medical Physics*, vol. 39, no. 9, pp. 5419–5428, 2012.
- [96] Yao Wang, "Graph-based segmentation of lymph nodes in ct data", Master's thesis, University of Iowa, 2010.
- [97] D. L. Hill, D. J. Hawkes, M. J. Gleeson, T. C. Cox, A. J. Strong, W. L. Wong, C. F. Ruff,

- N. D. Kitchen, D. G. Thomas, and A. Sofat, "Accurate frameless registration of mr and ct images of the head: applications in planning surgery and radiation therapy.", *Radiology*, vol. 191, no. 2, pp. 447–454, May 1994.
- [98] C. Studholme, D. L. G. Hill, and D. J. Hawkes, "Automated 3d registration of truncated mr and ct images of the head", in *Proceedings of the 1995 British conference on Machine vision (Vol. 1)*, Surrey, UK, UK, 1995, BMVC '95, pp. 27–36, BMVA Press.
- [99] C. Studholme, D.L.G. Hill, and D.J. Hawkes, "Automated 3-d registration of mr and ct images of the head", *Medical Image Analysis*, vol. 1, no. 2, pp. 163 – 175, 1996.
- [100] Petra A. van den Elsen, Evert-Jan D. Pol, Thilaka S. Sumanaweera, Paul F. Hemler, Sandy Napel, and John R. Adler, "Grey value correlation techniques used for automatic matching of ct and mr brain and spine images", 1994, vol. 2359, pp. 227–237, SPIE.
- [101] R. P. Woods, J. C. Mazziotta, and S. R. Cherry, "Mri-pet registration with automated algorithm", *Journal on Computer Assisted Tomography*, , no. 17, pp. 536–546, 1993.
- [102] C. Studholme, "An overlap invariant entropy measure of 3d medical image alignment", *Pattern Recognition*, vol. 32, no. 1, pp. 71–86, Jan. 1999.
- [103] G. P. Penney, J. Weese, J. A. Little, P. Desmedt, D. L. G. Hill, and D. J. Hawkes, "A comparison of similarity measures for use in 2-d-3-d medical image registration", *IEEE Transactions on Medical Imaging*, vol. 17, no. 4, pp. 586–595, 1998.
- [104] Murali Rao, Y. Chen, B. C. Vemuri, and Fei Wang, "Cumulative residual entropy: a new measure of information", *IEEE Transactions on Information Theory*, vol. 50, no. 6, pp. 1220–1228, 2004.
- [105] Fei Wang and Baba C. Vemuri, "Non-rigid multi-modal image registration using cross-cumulative residual entropy", *International Journal of Computer Vision*, vol. 74, no. 2, pp. 201–215, August 2007.
- [106] Christian Hollensen, Mads Fogtmann Hansen, Liselotte Højgaard, Lena Specht, and Rasmus Larsen, "Segmenting the parotid gland using registration and level set methods", in *MICCAI Workshop Medical Image Analysis for the Clinic - A Grand Challenge*, Beijing, China, September 2010, pp. 305 – 312.
- [107] Liliane Ramus and Grégoire Malandain, "Multi-atlas based segmentation: Application to the head and neck region for radiotherapy planning", in *MICCAI Workshop Medical Image Analysis for the Clinic - A Grand Challenge*, Beijing, China, September 2010, pp. 281 – 288.
- [108] Vincent Garcia, Olivier Commowick, and Grégoire Malandain, "A robust and efficient block matching framework for non linear registration of thoracic CT images", in *Proceedings of A Grand Challenge on Pulmonary Image Registration (EMPIRE'10), held in conjunction with MICCAI'10*, September 2010.
- [109] David Gering and Ben Kalinosky, "Automatic segmentation of the parotid glands by situated bayesian classification", in *MICCAI Workshop Medical Image Analysis for the Clinic - A Grand Challenge*, Beijing, China, September 2010, pp. 289 – 296.
- [110] Jinzhong Yang, Yongbin Zhang, Lifei Zhang, and Lei Dong, "Automatic segmentation of parotids from ct scans using multiple atlases", in *MICCAI Workshop Medical Image Analysis*

- for the Clinic - A Grand Challenge, Beijing, China, September 2010, pp. 323 – 330.
- [111] Y. Papastavrou, D. Cash, D. Hawkes, and B. Hutton, "A multi-component similarity measure for improved robustness of non-rigid registration of combined fdg pet-ct head and neck images", in *4th European Conference of the International Federation for Medical and Biological Engineering*, Jos Sloten, Pascal Verdonck, Marc Nyssen, Jens Hauelsen, and Ratko Magjarevic, Eds., vol. 22 of *IFMBE Proceedings*, pp. 433–435. Springer Berlin Heidelberg, 2009.
  - [112] Xiao Han, Mischa S. Hoogeman, Peter C. Levendag, Lyndon S. Hibbard, David N. Teguh, Peter Voet, Andrew C. Cowen, and Theresa K. Wolf, "Atlas-based auto-segmentation of head and neck ct images", in *Proceedings of the 11th International Conference on Medical Image Computing and Computer-Assisted Intervention, Part II*, Berlin, Heidelberg, 2008, MICCAI '08, pp. 434–441, Springer-Verlag.
  - [113] P. Cachier and X. Pennec, "3d non-rigid registration by gradient descent on a gaussian-windowed similarity measure using convolutions", in *Proc. IEEE Workshop Mathematical Methods in Biomedical Image Analysis*, 2000, pp. 182–189.
  - [114] D. Mattes, D. R. Haynor, H. Vesselle, T. K. Lewellen, and W. Eubank, "Pet-ct image registration in the chest using free-form deformations", *IEEE Transactions on Medical Imaging*, vol. 22, no. 1, pp. 120–128, 2003.
  - [115] A. du Bois d'Aische, M. de Craene, B. Macq, and S. K. Warfield, "An articulated registration method", in *Proc. IEEE Int. Conf. Image Processing ICIP 2005*, 2005, vol. 1.
  - [116] A. du Bois d'Aische, M. De Craene, B. Macq, and S. K. Warfield, "An improved articulated registration method for neck images", in *Proc. 27th Annual Int. Conf. of the Engineering in Medicine and Biology Society IEEE-EMBS 2005*, 2005, pp. 7668–7671.
  - [117] Aloys du Bois d'Aische, Mathieu De Craene, Xavier Geets, Vincent Gregoire, Benoit Macq, and Simon K. Warfield, "Estimation of the deformations induced by articulated bodies: Registration of the spinal column", *Biomedical Signal Processing and Control*, vol. 2, no. 1, pp. 16 – 24, 2007.
  - [118] M. Urschler, H. Ditt, and H. Bischof, "Partially rigid bone registration in ct angiography", in *Proc of the Computer Vision Winter Workshop (CVWW)*, Telc, Czech Republic, Feb. 2006, pp. 34–39.
  - [119] Adil Al-Mayah, Joanne Moseley, Shannon Hunter, Mike Velec, Lily Chau, Stephen Breen, and Kristy Brock, "Biomechanical-based image registration for head and neck radiation treatment", *Physics in Medicine and Biology*, vol. 55, no. 21, pp. 6491, 2010.
  - [120] R.H. Huesman, G.J. Klein, J.A. Kimdon, C. Kuo, and S. Majumdar, "Deformable registration of multimodal data including rigid structures", *IEEE Transactions on Nuclear Science*, vol. 50, no. 3, pp. 389 – 392, june 2003.
  - [121] Yangqiu Hu and David R. Haynor, "Multirigid registration of mr and ct images of the cervical spine", 2004, vol. 5370, pp. 1527–1538, SPIE.
  - [122] Peter Čech, Adrian Andronache, Liping Wang, Gábor Székely, and Philippe Cattin, "Piecewise rigid multimodal spine registration", in *Bildverarbeitung für die Medizin 2006*, Heinz Handels, Jan Ehrhardt, Alexander Horsch, Hans-Peter Meinzer, and Thomas Tolxdorff, Eds., Informatik



- aktuell, pp. 211–215. Springer-Verlag, 2006.
- [123] Wei Li, Lixu Gu, Jianrong Xu, and Kan Cheng, “Registration of the vertebrae”, in *Proc. Int. Conf. Information Technology and Applications in Biomedicine ITAB 2008*, 2008, pp. 66–69.
- [124] Xia Li, Todd Peterson, John Gore, and Benoit Dawant, “Automatic inter-subject registration of whole body images”, in *Biomedical Image Registration*, Josien Pluim, Botjan Likar, and Frans Gerritsen, Eds., vol. 4057 of *Lecture Notes in Computer Science*, pp. 18–25. Springer Berlin / Heidelberg, 2006.
- [125] M. Baiker, J. Milles, A. M. Vossepoel, I. Que, E. L. Kaijzel, C. W. G. M. Lowik, J. H. C. Reiber, J. Dijkstra, and B. P. F. Lelieveldt, “Fully automated whole-body registration in mice using an articulated skeleton atlas”, in *Proc. 4th IEEE Int. Symp. Biomedical Imaging: From Nano to Macro ISBI 2007*, 2007, pp. 728–731.
- [126] Xenophon Papademetris, Donald Dione, Lawrence Dobrucki, Lawrence Staib, and Albert Sinusas, “Articulated rigid registration for serial lower-limb mouse imaging”, in *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2005*, James Duncan and Guido Gerig, Eds., vol. 3750 of *Lecture Notes in Computer Science*, pp. 919–926. Springer Berlin / Heidelberg, 2005.
- [127] Jung W. Suh, Dustin Scheinost, Donald P. Dione, Lawrence W. Dobrucki, Albert J. Sinusas, and Xenophon Papademetris, “A non-rigid registration method for serial lower extremity hybrid spect/ct imaging”, *Medical Image Analysis*, vol. 15, no. 1, pp. 96 – 111, 2011.
- [128] Miguel Á. Martín-Fernández, Rubén Cárdenes, Emma Muñoz Moreno, Rodrigo de Luis-García, Marcos Martín-Fernández, and Carlos Alberola-López, “Automatic articulated registration of hand radiographs”, *Image Vision Computing*, vol. 27, pp. 1207–1222, July 2009.
- [129] Haili Chui and Anand Rangarajan, “A new point matching algorithm for non-rigid registration”, *Computer Vision and Image Understanding*, vol. 89, no. 2-3, pp. 114–141, Feb. 2003.
- [130] David T. Gering, Arya Nabavi, Ron Kikinis, W. Eric L. Grimson, Nobuhiko Hata, Peter Everett, Ferenc A. Jolesz, and William M. Wells, III, “An integrated visualization system for surgical planning and guidance using image fusion and interventional imaging”, in *Proceedings of the Second International Conference on Medical Image Computing and Computer-Assisted Intervention*, London, UK, UK, 1999, MICCAI '99, pp. 809–819, Springer-Verlag.
- [131] Martin Baiker, Julien Milles, Jouke Dijkstra, Tobias D. Henning, Axel W. Weber, Ivo Que, Eric L. Kaijzel, Clemens W.G.M. Löwik, Johan H.C. Reiber, and Boudewijn P.F. Lelieveldt, “Atlas-based whole-body segmentation of mice from low-contrast micro-ct data”, *Medical Image Analysis*, vol. 14, no. 6, pp. 723 – 737, 2010.
- [132] P. J. Besl and H. D. McKay, “A method for registration of 3-d shapes”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 239–256, 1992.
- [133] Tobias Klinder, Robin Wolz, Cristian Lorenz, Astrid Franz, and Jörn Ostermann, “Spine segmentation using articulated shape models”, in *Medical Image Computing and Computer Assisted Intervention - MICCAI 2008*, sep 2008, vol. 0.
- [134] Jonathan Boisvert, Xavier Pennec, Hubert Labelle, Farida Cheriet, Nicholas Ayache, Sainte-justine Hospital, and Chemin De CÃ´te-ste catherine, “Principal spine shape deformation

- modes using riemannian geometry and articulated models", *Articulated Motion and Deformable Objects Proceedings*, vol. 4069, pp. 346–355, 2006.
- [135] Haili Chui, Lawrence Win, Robert Schultz, James S. Duncan, and Anand Rangarajan, "A unified non-rigid feature registration method for brain mapping", *Medical Image Analysis*, vol. 7, no. 2, pp. 113 – 130, 2003.
  - [136] James C. Spall, "Overview of the simultaneous perturbation method for efficient optimization", *Johns Hopkins APL Technical Digest*, vol. 19, no. 4, pp. 482–492, 1998.
  - [137] G. K. Rohde, A. Aldroubi, and B. M. Dawant, "The adaptive bases algorithm for intensity-based nonrigid image registration", *IEEE Transactions on Medical Imaging*, vol. 22, no. 11, pp. 1470–1479, 2003.
  - [138] K. K. Brock, M. B. Sharpe, L. A. Dawson, S. M. Kim, and D. A. Jaffray, "Accuracy of finite element model-based multi-organ deformable image registration", *Medical Physics*, vol. 32, no. 6, pp. 1647–1659, 2005.
  - [139] G. D. Forney, "The viterbi algorithm", *Proceedings of the IEEE*, vol. 61, no. 3, pp. 268–278, Mar. 1973.
  - [140] Marshall F. Tappen and William T. Freeman, "Comparison of graph cuts with belief propagation for stereo, using identical mrf parameters", in *ICCV'03*, 2003, pp. 900–907.
  - [141] Yuri Boykov and Vladimir Kolmogorov, "An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 9, pp. 1124–1137, September 2004.
  - [142] Lars Dornheim, Jana Dornheim, Heiko Seim, and Klaus D. Tönnies, K. D. nnies, "Aktive sensoren: Kontextbasierte filterung von merkmalen zur modellbasierten segmentierung.", in *Bildverarbeitung für die Medizin*, Heinz Handels, Jan Ehrhardt, Alexander Horsch, Hans-Peter Meinzer, and Thomas Tolxdorff, Eds. 2006, Informatik Aktuell, pp. 96–100, Springer.
  - [143] John Ashworth Nelder and Roger Mead, "A simplex method for function minimization", *Computer Journal*, vol. 7, pp. 308–313, 1965.
  - [144] T. L. Faber and E. M. Stokely, "Orientation of 3-d structures in medical images", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 10, no. 5, pp. 626–633, Sept. 1988.
  - [145] Matthias Kirschner, Meike Becker, and Stefan Wesarg, "3d active shape model segmentation with nonlinear shape priors", in *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2011*, Gabor Fichtinger, Anne Martel, and Terry Peters, Eds., vol. 6892 of *Lecture Notes in Computer Science*, pp. 492–499. Springer Berlin Heidelberg, 2011.
  - [146] Marc Alexa, "Linear combination of transformations", in *Proceedings of the 29th annual conference on Computer graphics and interactive techniques*, New York, NY, USA, 2002, SIGGRAPH '02, pp. 380–387, ACM.
  - [147] K. Pearson, "On lines and planes of closest fit to systems of points in space", *Philosophical Magazine*, vol. 2, no. 6, pp. 559–572, 1901.
  - [148] B. Moghaddam and A. Pentland, "Probabilistic visual learning for object representation", *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 19, no. 7, pp. 696 –710,

jul 1997.

- [149] R. M. Wilcox, "Exponential operators and parameter differentiation in quantum physics", *Journal of Mathematical Physics*, vol. 8, no. 4, pp. 962–982, 1967.
- [150] Bram Van Ginneken, Tobias Heimann, and Martin Styner, "3d segmentation in the clinic: A grand challenge", in *MICCAI Workshop on 3D Segmentation in the Clinic: A Grand Challenge.*, 2007.
- [151] Jay West, J. Michael Fitzpatrick, Matthew Y. Wang, Benoit M. Dawant, Calvin R. Maurer, Jr., Robert M. Kessler, Robert J. Maciunas, Christian Barillot, Didier Lemoine, Andre Collignon, Frederik Maes, Paul Suetens, Dirk Vandermeulen, Petra A. van den Elsen, Sandy Napel, Thilaka S. Sumanaweera, Beth Harkness, Paul F. Hemler, Derek L. G. Hill, David J. Hawkes, Colin Studholme, J. B. Antoine Maintz, Max A. Viergever, Gregoire Malandain, Xavier Pennec, Marilyn E. Noz, Gerald Q. Maguire, Jr., Michael Pollack, Charles A. Pelizzari, Richard A. Robb, Dennis Hanson, and Roger P. Woods, "Comparison and evaluation of retrospective intermodality brain image registration techniques", *JOURNAL OF COMPUTER ASSISTED TOMOGRAPHY*, vol. 21, pp. 554–566, 1997.
- [152] Ion P. I. Pappas, Martin Styner, Puja Malik, Luca Remonda, and Marco Caversaccio, "Automatic method to assess local ct-mr imaging registration accuracy on images of the head", *American Journal of Neuroradiology*, vol. 26, no. 1, pp. 137–144, 2005.
- [153] Ben Glocker, J. Feulner, Antonio Criminisi, D.R. Haynor, and E. Konukoglu, "Automatic localization and identification of vertebrae in arbitrary field-of-view ct scans", in *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2012*, Nicholas Ayache, Hervé Delingette, Polina Golland, and Kensaku Mori, Eds., vol. 7512 of *Lecture Notes in Computer Science*, pp. 590–598. Springer Berlin Heidelberg, 2012.
- [154] T. Rohlfing, "Image similarity and tissue overlaps as surrogates for image registration accuracy: Widely used but unreliable", *IEEE Transactions on Medical Imaging*, vol. 31, no. 2, pp. 153–163, feb. 2012.
- [155] S. Klein, M. Staring, K. Murphy, M. A. Viergever, and J. Pluim, "elastix: A toolbox for intensity-based medical image registration", *IEEE Transactions on Medical Imaging*, vol. 29, no. 1, pp. 196–205, 2010.
- [156] Joseph O'Rourke, "Finding minimal enclosing boxes", *Internat. J. Comput. Inform. Sci.*, vol. 14, pp. 183–199, 1985.
- [157] Ivo Wolf, Marcus Vetter, Ingmar Wegner, Thomas Böttger, Marco Nolden, Max Schäbinger, Mark Hastenteufel, Tobias Kunert, and Hans-Peter Meinzer, "The medical imaging interaction toolkit", *Medical Image Analysis*, vol. 9, no. 6, pp. 594 – 604, 2005.
- [158] Matthias Kirschner, Florian Jung, and Stefan Wesarg, "Automatic prostate segmentation in mr images with a probabilistic active shape model", in *PROMISE12 - MICCAI 2012 Grand Challenge on Prostate MR Image Segmentation*, Dean Barratt, Ed., 2012, pp. 28–35.
- [159] Wissam El Hakimi and Stefan Wesarg, "Confidence map based super-resolution reconstruction", pp. 831440–831440–10, 2012.
- [160] Guy Jerusalem, Roland Hustinx, Yves Beguin, and Georges Fillet, "Evaluation of therapy for

- lymphoma", *Seminars in Nuclear Medicine*, vol. 35, no. 3, pp. 186 – 196, 2005.
- [161] Teena Steger and Martin Hosbach, "Navigated bronchoscopy using intraoperative fluoroscopy and preoperative ct.", in *ISBI. 2012*, pp. 1220–1223, IEEE.
- [162] Teena Steger and Stefan Wesarg, "Marker detection evaluation by phantom and cadaver experiments for c-arm pose estimation fiducial", in *Medical Imaging 2013: Image Processing (to appear)*. 2013, SPIE.
- [163] M. Erdt, M. Kirschner, K. Drechsler, S. Wesarg, M. Hammon, and A. Cavallaro, "Automatic pancreas segmentation in contrast enhanced ct data using learned spatial anatomy and texture descriptors", in *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on*, 30 2011-april 2 2011, pp. 2076 –2082.

# Abkürzungsverzeichnis

|             |  |
|-------------|--|
| <b>AAM</b>  | Active Appearance Models, p. 20.                     |
| <b>ABA</b>  | Adaptive Bases Algorithm, p. 57.                     |
| <b>ADC</b>  | Apparent Diffusion Coefficient, p. 14.               |
| <b>AS</b>   | Automatisches Schwellwertverfahren, p. 30.           |
| <b>ASD</b>  | Average Surface Distance, p. 137.                    |
| <b>ASM</b>  | Active Shape Models, p. 20.                          |
| <b>BP</b>   | Belief Propagation, p. 70.                           |
| <b>CCRG</b> | Confidence Connected Region-Growing, p. 31.          |
| <b>CT</b>   | Computertomographie, p. 1.                           |
| <b>DFFS</b> | Distance From Feature Space, p. 54.                  |
| <b>DIFS</b> | Distance In Feature Space, p. 54.                    |
| <b>DOF</b>  | Degrees of Freedom, p. 22.                           |
| <b>DSC</b>  | Dice Similarity Coefficient, p. 31.                  |
| <b>DWI</b>  | Diffusion Weighted Imaging, p. 14.                   |
| <b>EBS</b>  | Elastic Body Splines, p. 23.                         |
| <b>ECC</b>  | Extended Confidence Connected Region-Growing, p. 31. |
| <b>EMSM</b> | Extended Mass Spring Model, p. 138.                  |
| <b>FDG</b>  | Fluordesoxyglucose, p. 185.                          |
| <b>FEM</b>  | Finite Element Method, p. 49.                        |
| <b>GA</b>   | Gradientenabstieg, p. 83.                            |
| <b>GC</b>   | Graph Cut, p. 42.                                    |
| <b>GCS</b>  | Graph Cut Segmentation, p. 31.                       |
| <b>HD</b>   | Hausdorff-Distanz, p. 33.                            |
| <b>HU</b>   | Hounsfield Units, p. 13.                             |
| <b>ICP</b>  | Iterative Closest Point, p. 52.                      |
| <b>ITK</b>  | Insight Toolkit, p. 155.                             |
| <b>LS</b>   | Lokales Schwellwertverfahren, p. 30.                 |

|               |  |
|---------------|--|
| <b>MI</b>     | Mutual Information.  |
| <b>MITK</b>   | Medical Imaging Interaction Toolkit, p. 183.                 |
| <b>MRF</b>    | Markov Random Field, p. 19.                                  |
| <b>MRT</b>    | Magnetresonanztomographie, p. 1.                             |
| <b>MSD</b>    | Maximum Surface Distance, p. 137.                            |
| <b>MSM</b>    | Mass Spring Model, p. 138.                                   |
| <b>PCA</b>    | Principal Component Analysis, p. 20.                         |
| <b>PET</b>    | Positron Emission Tomography, p. 1.                          |
| <b>PSAM</b>   | Parametric Shape and Appearance Model, p. 31.                |
| <b>RECIST</b> | Response Evaluation Criteria In Solid Tumours, p. 2.         |
| <b>ROI</b>    | Region of Interest, p. 22.                                   |
| <b>SMSM</b>   | Stable Mass Spring Model, p. 35.                             |
| <b>SN</b>     | Snakes, p. 44.   |
| <b>STAPLE</b> | Simultaneous Truth and Performance Level Estimation, p. 136. |
| <b>TNM</b>    | Tumor-Node-Metastasis, p. 2.                                 |
| <b>TPS</b>    | Thin Plate Splines, p. 23.                                   |
| <b>TRE</b>    | Target Registration Error, p. 49.                            |
| <b>USPI</b>   | Ultrasmall Superparamagnetic Iron Oxyde, p. 14.              |
| <b>VOE</b>    | Volumetric Overlap Error, p. 136.                            |
| <b>WT</b>     | Wasserscheidentransformation, p. 30.                         |

# Lebenslauf

## Persönliche Daten

Name: Sebastian Steger  
Geburtstag und Ort: 23. April 1983 in München  
Familienstand: verheiratet  
Nationalität: österreichisch

## Bildung

2007                      Universitätsabschluss zum Dipl.-Ing. der Technischen Informatik an der Technischen Universität Berlin, Diplomarbeitsthema: Embedded Trellis Quantization  
  
2002 – 2007            Studium der Technischen Informatik an der Technischen Universität Berlin, Berlin, Deutschland  
  
1989 – 2002            Grundschule und Gymnasium in Kirchheim b. München und Berlin, Deutschland

## Berufserfahrung

2009 –                      Wissenschaftlicher Mitarbeiter, Fraunhofer IGD, Darmstadt, Deutschland, Fokus: Segmentierung und Registrierung von medizinischen Bilddaten  
  
2008                      Softwareentwickler, Accusoft, Tampa, FL, USA, Fokus: Softwareentwicklung von Bildverarbeitungsroutinen  
  
2007                      Hilfwissenschaftler, Institut für Mathematik, Technische Universität Berlin, Berlin, Deutschland, Fokus: Forschung im Bereich der Bilddatenkompression  
  
2003 – 2006            Softwareentwickler, PACE Aerospace Engineering and Information Technology GmbH, Berlin, Deutschland, Fokus: Softwareentwicklung  
  
1999 – 2002            Freiberuflicher Softwareentwickler